# DOCUMENT RESUME

ED 176 248 CS 005 031

| | |
|---|---|
| AUTHOR | Greene, Frank P., Ed. |
| TITLE | Investigations Relating to Mature Reading; Twenty-first Yearbook of the National Reading Conference. Volume I. |
| INSTITUTION | National Reading Conference, Inc. |
| PUB DATE | 72 |
| NOTE | 338p.; For related document, see CS 005 032 ; Proceedings of the National Reading Conference (21st, Tampa, Florida, December 2-4, 1971) |
| EDRS PRICE | MF01/PC14 Plus Postage. |
| DESCRIPTORS | *Adult Reading Programs; Advance Organizers; Cloze Procedure; Elementary Secondary Education; Higher Education; Instructional Materials; Measurement Techniques; Reading Processes; *Reading Programs; *Reading Research; *Teacher Education; Teaching Techniques |
| IDENTIFIERS | *National Reading Conference |

ABSTRACT

This is one of two volumes of papers drawn from the twenty-first annual meeting of the National Reading Conference held in 1971. Focusing on research and teacher training concerns relating to mature reading, the 43 papers are arranged into the following categories: advanced organizers, cloze investigations, langauge investigations, measurement problems, Right to Read reports, teaching materials and approaches, and teacher training programs. An annual review of research is also presented. (FL)

***********************************************************************
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
***********************************************************************

ED176248

# INVESTIGATIONS RELATING TO MATURE READING

## TWENTY-FIRST YEARBOOK OF THE NATIONAL READING CONFERENCE

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Edited by
FRANK P. GREENE
McGill University

Published by
The National Reading Conference, Inc.
Milwaukee, Wisconsin 53233

1972

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
The National Reading
Conference, Inc.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST AVAILABLE COPY

CS005031

OFFICERS AND BOARD OF DIRECTORS
OF THE
NATIONAL READING CONFERENCE, INC.

President: Wendell W. Weaver, University of Georgia
Vice-President: Earl F. Rankin, University of Kentucky
Treasurer: L. L. Van Gilder, Marquette University
Secretary: Hazel Horn Carroll, Southern Methodist University
Past President: Alton L. Raygor, University of Minnesota
Edward B. Fry, Rutgers—The State University
Alvin J. Lowe, University of South Florida
Bernard Schmidt, Purdue University
Editor of the Yearbook, Frank P. Greene, McGill University

Copyright 1972
The National Reading Conference, Inc.
All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

Distributed by
NATIONAL READING CONFERENCE, INC.
Reading Center, Marquette University
Milwaukee, Wisconsin 53233

The ideas expressed in these papers are those of the authors and do not necessarily represent the viewpoints of the National Reading Conference

Printed in the United States of America

$5.00

# PREFACE

Remarks made in the preface to last year's Yearbook need to be corrected. Legal majority now comes at 18, so NRC did not need to wait until this 21st conference to become of age. The question of maturity remains. It has been suggested by some that due to the change in the age of majority, NRC may be almost over the hill. Others have suggested that reality lies in the other direction and that NRC simply enjoys a continued adolesence. However, except for decent intervals during the discussions of vital issues when a second childhood seems to predominate, NRC is in fact most mature.

As a conference, maturity may be noted in the multiple paper presentations made at multiple sessions and in the increased active membership. As a Conference, maturity may be noted in the many reports which exhibit both delayed gratification (long term studies) and high level analytic cognition (sharp problem statements and clear data manipulations). It is reasonable to expect that such mature behavior will be able to provide some acceptable answers to some of the long standing problems of college and adult reading.

This Yearbook contains papers from the annual conference held December 2, 3 and 4 December, 1971 at Tampa, Florida. The papers selected for inclusion in the Yearbook relate to the research and to the teacher training concerns of the membership.

In addition to this Yearbook, NRC is also publishing *College reading: problems and programs of junior and senior colleges.* This book of readings was also constructed from papers presented at the 21st annual conference. Papers selected for inclusion in *College reading* relate to setting up, operating, and evaluating reading and study skills programs at two and four year colleges.

In this 21st Yearbook, a memorium to Homer Carter, longtime member and friend of NRC, appears as the first paper. The annual review of research on college-adult reading appears as the last paper. The other papers have been grouped into topics and appear within these categories as indicated in the table of contents.

Happy reading,
F. P. G.

# TABLE OF CONTENTS

Preface ........ iii

Memorium to Homer L. J. Carter ........ 1

Advanced Organizer Symposium

Reception learning and advanced organizers: implications for reading research
*B. Elgit Blanton* ........ 3

The effects of advanced organizers and grade level upon the reception, learning and retention of general science content
*Richard F. Barron* ........ 8

The effect of advanced organizers upon meaningful reception, learning and retention of social studies content
*Thomas H. Estes* ........ 16

The effects of advanced organizers in reading for the retention of specific facts
*Bob W. Jerrolds* ........ 22

Cloze Investigations

Development of telegraphic prose based on a random word deletion scheme
*Clessen J. Martin and Mary Anne Herndon* ........ 30

Cloze and closure: a second analysis
*Fred W. Ohnmacht and James T. Fleming* ........ 35

Validation of the cloze procedure as a measure of readibility with high school, trade school, and college populations.
*Joe Peterson, Nat Peters and Ed Paradis* ........ 45

A comparison of the reading performance of college students with conventional versus negative internal (intra-article) reading flexibility
*Earl F. Rankin and Tom J. Kehle* ........ 51

The adequacy of the cloze in measuring comprehension of different logical patterns
*Ned H. Ratekin* ........ 59

Unilateral versus bilateral cloze performance
*Wilson L. Taylor* ........ 66

Speculation on cloze as a search procedure
*J. Jaap Tuinman* ........ 74

Language Investigations

Levels of syntactic realization in oral reading
*Eric Brown* ........ 85

Experiments in children's perception of words and word boundaries
*Albert J. Kingston, Wendell W. Weaver and Leslie E. Figa* 91

Paradigmatic language training as a method of increasing reading achievement
*George McNinch* 100

A comparison study of semantic and syntactic cueing by low reading performance college freshmen.
*Allan Roy Ohaver* 110

Reading, understanding, and conceptual situations
*Ernst von Glasersfeld* 119

Measurement Problems

Measuring change—a time dimensional problem
*Nicholas J. Anastasiow* 128

A comparison of psychometric models used in measuring change
*Fred B. Davis* 134

Effects of typeface on legibility, pleasingness, comprehension, and rate
*John Follman, A. J. Lowe, Edward Uprichard, Loren Roberts, Melvin Velleme* 141

Measuring change: some problems and solutions
*James L. Wardrop* 146

Heart rate and reading
*David M. Wark, Diana Tostenrud and George Nelms* 153

Aptitude interaction analysis: a necessity in reading research
*John A. Wood and Anne Marie Bernazza Haase* 161

Right to Read Reports

Developmental processes and reading processes
*Irene Athey* 171

Modeling the reading process: promise and problems
*John J. Geyer* 183

A critical review of the information base for current assumptions regarding the status of instruction and achievement in reading in the United States
*Patricia Elias* 191

Results of the survey of the literature on methods and materials in reading
*Martha J. Maxwell* 203

Teaching Materials and Approaches

Design concepts for contingency management of delinquent adolescents
*Robert L. Aaron* .................... 212

Peer versus teacher correction of classwork and selected reading criteria
*Joan S. Grice and Lee R. Wolfe* .................... 218

Reading by patterns: a psycholinguistic remedial tutorial program for young adults
*Carl A. Lefevre* .................... 228

Individualized instructional objectives in reading methods courses
*Patricia R. McClendon* .................... 237

The Wisconsin design for reading skill development

Rational and overview
*Wayne Otto* .................... 245

The word attack segment
*Eunice N. Askov* .................... 249

The study skill component
*Karlyn Kamm* .................... 253

The comprehension component
*Robert D. Chester* .................... 258

Implementation and field testing
*Robert T. Rude* .................... 265

Teacher Training Programs

Direct experience reading tasks

Training reading teachers
*Edward B. Fry* .................... 273

Psychological development
*Josephene Goldsmith* .................... 275

Syntax, word part deletions and linguistic terminology
*Joseph V. DePierro* .................... 277

Reading measurement
*Mary M. Byrd* .................... 279

Comprehension and study skills
*Joseph P. Sikoryak* .................... 281

Reading content specialists for the junior high school
*Harry Singer* .................... 283

Model of the dynamic activity of reading and the reading teacher
*Edmund Zazzera* ........ 292

Annual Review of Research
*Emery P. Bliesmer* ........ 298

**HOMER L. J. CARTER**
**1894–1971**

Professor Homer L. J. Carter, after an illness of three months, died on January 11, 1971. He was 76.

A graduate of Wayne State University, he earned his Master's degree from Ohio State University in 1931. His professional experience spanned a 50-year period. He began his teaching career in a rural school in Michigan. Later he served as a teacher, principal, and elementary supervisor in the Royal Oak Public Schools. He joined the faculty of Western Michigan University in 1928 as Associate Director of Research. At the time of his retirement from Western he held the rank of Professor and was Director of the Psycho-Educational Clinic.

Professor Carter established the Psycho-Educational Clinic, now known as the Reading Center and Clinic, in 1932. In 1944 he developed a reading laboratory for college students, and during the same year he initiated a series of reading demonstrations and discussions as an integral part of the course, Educational Therapy in Reading. The first council in the State of Michigan of the International Reading Association grew out of one of his summer classes of Educational Therapy in Reading. In 1960 Professor Carter stimulated the group to join with the Reading Center and Clinic in publishing *Reading Horizons*. He served as editor of this quarterly until his death. In

1964, at the time of his retirement, the chapter honored him by naming it the Homer L. J. Carter Reading Council.

Professor Carter was the senior author of seven books in the area of reading and college adjustment. The first was published in 1949 and the last, after his death, in 1971. Throughout the years he was actively engaged in research and the author of over 100 articles which have appeared in professional journals.

Professor Carter served as president of the Michigan Reading Association, the Michigan Psychological Association, and the Michigan Academy of Science, Arts and Letters. In 1951 he was signally honored by being elected a fellow of the American Association for the Advancement of Science. He considered it a privilege to be an active member of such professional organizations as the American Psychological Association, the International Reading Association, the College Reading Association, and the National Reading Conference. He was particularly proud of his affiliation with the National Reading Conference and was one of its most loyal supporters.

Homer L. J. Carter was a man of strong convictions. He believed that service to others is man's greatest gift to the world. He believed that every person should set up worthwhile goals and should work earnestly and diligently to achieve these objectives. He believed that there is no greater satisfaction in life than a job well done. He was a man who lived his beliefs.

Professor Carter was a forceful writer and an excellent teacher who gave much of himself in his constant efforts to help others and to make a worthwhile contribution in the field of reading. He inspired many of his students to follow in his footsteps and had a stimulating influence upon all who had the good fortune to work closely with him. It will be impossible for many of us to participate in the field of reading without remembering the enthusiasm and spirited teaching of Homer L. J. Carter.

B. ELGIT BLANTON
Indiana University

# RECEPTION LEARNING AND ADVANCED ORGANIZERS: IMPLICATIONS FOR READING RESEARCH

Perhaps the most common definition of comprehension found in the literature is the non-operational difinition in terms of skills: Investigators have merely listed the skills and abilities believed involved in comprehension. These skills range from identifying, organizing, interpreting and evaluating ideas, (Staiger and Bliesmer, 1965), to assimilation and accomodation (Stauffer, 1967). This concept about comprehension appears to be widely accepted among reading specialists. Unfortunately, when an attempt is made to give substantive content to this kind of difinition, the problems are enormous. The substantive content which the teacher of reading needs is only beneficial in the form of specific activities which he can manifest and in turn impose on the learner to bring about changes in reading behavior. In short, if changes in behavior are designated as identifying, organizing, interpreting, and evaluating ideas, it is necessary that the reading teacher know the particular manipulations which will culminate in these high-level constructs.

In view of the fact that learning from written materials is a major part of the education curriculum, the paucity of specific manipulations which may be imposed on the learner to facilitate learning by reading is a major educational problem. Researchers have noted, however, that a number of operations related to, but apart from, the reading act have a facilitative effect on learning by reading. This review will focus on researchers and their collaborators, who have carried on programmatic research on this topic. Interestingly enough, none of the studies reviewed have been considered reading studies either by the researchers or by their audience.

### AUSUBEL'S THEORY OF RECEPTION LEARNING

David P. Ausubel is a cognitive psychologist and the studies reviewed here are a part of a larger group of studies attempting to make theoretical points about verbal learning. The findings of these studies, nevertheless, have implications for learning by reading.

Ausubel's (1963) theory of meaningful reception learning postulates the existence of a cognitive structure that is hierarchically organized in terms of highly inclusive conceptual traces under which are subsumed less inclusive concepts. Ausubel further asserts that existing cognitive structure is the primary factor affecting meaningful verbal learning and retention.

In the theory of meaningful verbal learning a major concept, is that new material which is to be learned must be related to the learner's existing cognitive structure. In other words, new subject matter is acquired as unfamiliar knowledge is integrated, or subsumed, into the learner's hierarchical cognitive structure. Subsumption, then, is the process of incorporating new subject matter into a larger, comprehensive knowledge system.

As might be expected, Ausubel contends that the forgetting of materials is a process of obliterative subsumption. In explaining his theory of forgetting, Ausubel writes:

> Although the stability of meaningful material is initially enhanced by anchorage to relevant conceptual foci in the learner's cognitive structure, such material is gradually subjected to the erosive influence of the conceptualizing trend in cognitive organization. Because it is more economical and less burdensome to retain a single inclusive concept than to remember a large number of more specific items, the import of the latter tends to be incorporated by the generalized meaning of the former. When this second or obliterative stage of subsumption begins, the specific items become progressively less dissociable as entities in their own right until they are no longer available and are said to be forgotten (1962, p. 217.).

In short, forgetting is a process whereby more specific items of information are subsumed under a single or more inclusive concept.

From the above discussion of Ausubel's theoretical formulations, it is clear that if a learner is to acquire new concepts from verbal material, he must have an existing cognitive structure upon which new concepts can be subsumed or anchored. On the other hand, if the specific cognitive elements are not available for anchorage of specific new input, subsequent achievement is affected.

## ADVANCE ORGANIZERS

In order to facilitate the process of acquisition and retention of meaningful material, Ausubel advocates the use of advance organizers:

> . . . the most efficient way of facilitating retention is to introduce appropriate subsumers and make them part of the cognitive structure prior to the actual presentation of the learning task. The introduced subsumers thus become advance "Organizers" or anchoring foci for the reception of new material. In effect, they provide an introductory overview at the appropriate level of conceptualization (1963 p. 29).

Ausubel goes on to identify three important attributes of advance organizers that support their incorporation into learning materials:

1. Specifically constructed advance organizers based on specific learning materials are an efficient and appropriate way to facilitate the subsumption of subject matter. The deliberate introduction of relevant organizers into the learning cognitive structure is an effective method of manipulating the relevant attributes of cognitive structure in order to maximize the subsumption of more detailed subject matter.
2. Organizers are more effective for learning informational data and factual material than for the learning of abstract materials (concepts, principles, and generalizations) because abstractions are inherently hierarchical and contain their own inclusive organizers.
3. The effect of advance organizers will be more obvious if the learning material is structured to proceed from highly inclusive and theoretical subject matter to less inclusive and more specific learning material. Moreover, Ausubel postulates that organizers presented in advance of the learning material will facilitate conceptual integration more effectively than organizers presented concurrently with the learning material.

At least two kinds of organizers are defined by Ausubel. When the learner is to be exposed to unfamiliar material, an expository organizer is used to provide subsumers. In the case of relatively familiar material, a comparative organizer is used to enable the learner to discriminate between new and existing concepts which are different but confusable.

Before moving on, it seems appropriate to point out that Ausubel differentiates between advance organizers and overviews usually found in textbooks. From his point of view, overviews are generally written at the same level of abstraction and generality as the materials to be learned. Moreover, overviews seem to accomplish their effects through repetition and emphasizing key words and concepts. Organizers, on the other hand, are written at a higher level of abstraction and generality than the material to be learned, and provide relevant subsuming concepts. Parenthetically, it might be added that nowhere does Ausubel define the term advance organizer operationally. Still another shortcoming is the fact that no procedure for generating an advance organizer is offered in the literature.

Briefly, then, the advance organizer is a set of material designed specifiically to facilitate the incorporation of a given body of knowledge into the cognitive structure of the learner. Of the advance organizer, Ausubel (1963) says that it must be written on a higher level of abstraction, inclusiveness, and generality than that of the new material to be presented. Once the advance organizer has been established in the learner's cognitive structure, the new ideas can be subsumed under its more inclusive, abstract, and general propositions.

## RELATED RESEARCH

As mentioned earlier, the task which Ausubel imposes is to read passages of material on specific content. The experimental variable is an advance organizer—another, usually shorter, passage of reading material—which may or may not be related to the material to be learned. The dependent variables are multiple-choice achievement tests constructed to determine the degree to which learning of material has taken place.

In his first study, Ausubel (1960) tested the hypothesis that the learning and retention of unfamiliar but meaningful verbal material can be facilitated by an advance organizer. Eighty undergraduates were assigned to an organizer or no organizer group prior to reading a 2500-word passage ont he metalurgical properties of steel. An analysis of the data confirmed the hypothesis. Generalizing from the results of the experiment, however, is difficult for three reasons. First, test reliabilities were not reported. Second, it is difficult to determine whether the test is measured substantive or verbatum learning. Finally, exactly how the advance organizer was generated is not clear.

The above study was followed by an attempt to compare the facilitative effects of a comparative organizer and an expository organizer (Ausubel and Fitzgerald, 1961). Undergraduate subjects were assigned to organizer or no organizer conditions prior to reading a passage on Buddhism. A test administered after a ten-day intervel revealed that both organizers were significantly effective. Unfortunately, it is impossible to make generalizations from this study for the same reasons mentioned above.

To summarize, a review of the research reveals that studies exploring the facilitative effects of advance organizers have been reported by educational psychologists and reading specialists. For a number of reasons, however, only limited generalization can be made from the research reviewed. In additiont o not operationally defining the term advance organizer, researchers have failed to define procedures for generating advance organizers. The rather loose procedures reported have resulted in a potpourri of advance organizers. Thus, the position for advance organizers looses its potency.

The above argument notwithstanding, it will be remembered that investigators have neglected to report statistics for the criterion measures used in their experiments. This is an important point in light of the fact that most measures used have been approximately six to fourteen items long. Moreover, it is difficult to determine if the measures used were reading dependent. In other words, it is difficult to determine the effects of treatment on learning by reading if evidence demonstrating that learners did not possess the information of interest prior to reading the material is lacking.

## IMPLICATIONS FOR READING RESEARCH

If additional evidence indicates that Ausubel's theory is acceptable, reading specialists might consider the following points. First, written materials used in the schools should be rewritten with concern for cognitive structuring. Similarly, most sets of materials should have advance organizers written for them. Third, it is possible that having curriculum materials dependent upon specific prior materials would freeze the student in an increasingly inflexible program. Fourth, diagnostic procedures and instruments would have to be developed in order to determine each learner's cognitive structure and his subsequent placement in the curriculum. Finally, it would be mandatory that one have an intimate knowledge of the background and mental operations of each learner who is to read material before organizers could be constructed for the learner: Individual differences in cognitive structure would require that advance organizers be written for every learner or groups of learners for each set of materials used in the curriculum.

## REFERENCES

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1960, 51, 267-272.

Ausubel, D. P. A subsumption theory of meaningful verbal learning and retention. *The Journal of General Psychology*, 1962, *66*, 213-221.

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune and Stratton, 1963.

Ausubel, D. P. & Fitzgerald, D. The role of discriminability in meaningful verbal learning and retention. *Journal of Educational Psychology*, 1961, *52*, 266-274.

Staiger, R. C. & Bliesmer, E. P. Reading comprehension in the high school. *Education*, 1956, *76*, 563-567.

Stauffer, R. G. Reading as a cognitive process. *Elementary English*, 1967, *44*, 343-348..

RICHARD F. BARRON
Oakland University

# EFFECTS OF ADVANCE ORGANIZERS AND GRADE LEVEL UPON THE RECEPTION LEARNING AND RETENTION OF GENERAL SCIENCE CONTENT[1] [2]

This study attempted to determine the effects of graphic and prose advance organizers at each of seven grade levels, six through twelve.

## BACKGROUND

David Ausubel (1963, 1968) has developed an educational psychology that is intimately related to learning as it typically occurs in schools. At the core of this psychology is the proposition that a principal variable influencing new learning in a subject matter field is one's existing background of knowledge, or cognitive structure. Ausubel maintains that new meanings in any discipline are acquired only in relation to a previously learned background of relevant concepts and principles. Thus, if existing cognitive structure is clear, stable and organized, new learning will be enhanced. Conversely, if existing cognitive structure is unclear and disorganized, new learning will be impeded. Following this line of reasoning, Ausubel has argued that new learning and retention can be facilitated by strengthening relevant aspects of a learner's existing cognitive structure.

Ausubel (1960) has proposed the use of advance organizers as one means of strengthening existing cognitive structure. As developed by Ausubel, advance organizers are introductory prose passages written at a higher level of generality, abstraction, and inclusiveness than the actual learning material. These devices, "provide ideational scaffolding for the stable incorporation of more differentiated material in the learning passage." and "they increase discriminability between the new material and similar or ostensibly conflicting ideas in cognitive structure [1968, p. 148]."

---

[1] A note of thanks is due to Heather Campbell, Ellsworth Berget, and John Carney confreres from Syracuse University and to the English Department of the Tully Central Schools, Tully, N.Y., for their assistance in the conduct of this study.

[2] The research reported herein was performed pursuant to a grant with the Office of Education U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Although Ausubel's thinking is logically compelling, its implementation and evaluation have been beset by a number of problems. First, advance organizers have not been accorded an operational definition. Clearly, what is sufficiently general, abstract, and inclusive varies from one situation to another. Thus, it has been difficult for teachers and researchers to know whether particular organizers were appropriate for their intended audience.

Second, in studying the effects of advance organizers, previous investigators have tended to concentrate upon basic rather than applied research. Although the results of these studies are theoretically interesting, the points at which advance organizers achieve pratical utility have not been identified. Rather, previous research has consistently demonstrated the equivocal effects of a poorly defined treatment in a number of specific situations. Table 1 presents a summary of prior research.

Table 1
Previous Investigations of Advance Organizers

| Populations | Results | |
|---|---|---|
| | Supportive | Equivocal and/or Negative |
| Elementary | | Carter, Proger, Mann, & Haughley, 1970<br>Proger, Mann, Taylor, Morris, & Reckless, 1970. |
| Junior High School | | Triezenberg, 1968 |
| High School | Fitzgerald & Ausubel, 1963<br>Estes, Mills & Barron, 1969 | Jerrolds, 1967<br>Smith & Hess, 1969<br>Thelen, 1970<br>Proger, Taylor, Mann, Coalson, & Bayuk, 1970 |
| College | Ausubel, 1960<br>Ausubel & Fitzgerald, 1962<br>Murphy, 1962<br>Ausubel & Fitzgerald, 1962<br>Ausubel & Youssef, 1963<br>Scandura & Wells, 1967 | Baumman & Glass, 1969 |
| Adult | Grotelueshen & Sjogren, 1968 | |

The present investigation attempted to deal with the preceding problems by: developing an operationally defined advance organizer designed to meet general conditions specified by Ausubel; and testing the effects of this organizer relative to those of a traditional prose organizer in a strong inference research design.

## GRAPHIC ADVANCE ORGANIZERS

Graphic organizers were first suggested under the rubic of structured overviews (Barron, 1969). They have been defined as "visual and verbal presentations of the key vocabulary in a new learning task in relation to subsuming and/or parallel terms that presumably have previously been incorparated into the learner's cognitive structure [Estes, Mills & Barron, 1969, p. 41]." Graphic organizers have been based upon the same theoretical rationale as Ausubel's prose organizers and may be regarded as a "special form of advance organizers [Ausubel, 1969]."

In contrast to prose organizers, graphic organizers appear to hold a number of advantages. First, unlike prose organizers, graphic organizers have been operationally defined. It has been ascertained that they can be constructed and used by combining a minimum of training with a relatively simple set of directions:

1. Analyze the vocabulary of the learning task and list all the words that you feel are important for the student to understand.
2. Arrange the list of words until you have a scheme which depicts the interrelationships among the concepts particular to the learning task.
3. Add to the scheme vocabulary terms which you believe are understood by the students in order to depict relationships between the learning task and the discipline as a whole.
4. Evaluate the organizer. Have you clearly depicted major relationships? Can the overview be simplified and still effectively communicate the idea you consider to be crucial?
5. Introduce the students to the learning task by displaying the scheme and informing them why you arranged the terms as you did. Encourage them to contribute as much information as possible.
6. During the course of the learning task, relate new information to the organizer as it seems appropriate.

Second, whereas prose organizers are designed to be read by learners, graphic organizers call for an interaction between teacher and student. Thus, when using the latter device, a teacher is able to evaluate its appropriateness in relation to the learners' existing background of knowledge.

## STRONG INFERENCE

Platt (1964), among others, has called for the application of research procedures used in fast moving fields to educational studies. One of these techniques has been termed strong inference. Rather than continually measuring, defining, computing, and analyzing the same theoretical ground, the researcher attempts to refine crucial experiments aimed at disproving his hypotheses. While the overall utility of this approach in educational research is open to question, it can be valuable within certain limits. By developing a series of studies in terms of gradually refined populations and control variables, the educational researcher is able to indicate the degree to which a particular theoretical orientation achieves practical utility.

The present study attempted to initiate the preceding in the following way. Ausubel (1968) has indicated that thre are two conditions under which an advance organizer fails to perform its intended function. The first occurs when the learner is so lacking in background relevant to the learning task that there is little, or nothing, in his cognitive structure to be organized. The second situation occurs when the learner is so well grounded in background relevant to the learning task that provision of the organizer becomes superfluous.

These ideas were tested in the present study by holding both the learning task and organizer treatments constant across seven grade levels and aiming these materials at the middle level. It was anticipated that if Ausubel's thinking was correct an S-shaped learning curve would appear across the several grade levels. That is, the organizers should have initially demonstrated little or no effect due to the weak and unstable cognitive structures of the younger subjects. This effect should have gradually increased to some maximum point. Finally, it should have dissipated with older subjects who presumably would have well organized cognitive structures in relation to the learning task.

If, on the other hand, Ausubbel's assertions were sometimes less than generally applicable, it was anticipated that the learning curve across grade levels would resemble a straight or slightly curved line. In other words, subject variability within grades would negate the potential benefits of the organizer treatments.

## PROCEDURES

### *Subjects*

The study was conducted with students enrolled in grade six through twelve from the central New York State area. The school draws its students from rural and suburban environments. Excluded from the study were pupils classified as emotionally disturbed or mentally retarded, and those who were not in attendance during treatment implementation.

### *Design and statistical analysis*

The basic design for the study was a posttest-only control group design according to Campbell and Stanley (1963). Although the schematic representative of this design (see table 2) appears to be factorial test of significance for overall main effects and interactions were not of direct interest. Rather, the intent of the study was to explore potential specific interactions (in a non-factorial sense) between methods of cognitive organization and grade level.

In order to accomplish this purpose, two orthogonal planned comparison were posited at each grade level. The first comparison contrasted the average effects of the combined organizer treatments with those of the control. The second compared the effects of the two organizer treatments.

### *Learning passage*

The learning passage (adapted from Deitz, 1968) was selected primarily on the basis of its relevance to the regular eighth grade science curriculum. It was approximately 2300 words in length and it dealt with the characteristics of stars. The readability level of this passage was eighth grade level as determined by the Flesch Formula.

### *Organizers*

The graphic and prose organizers used in the study: reviewed terminology used in making comparisons; related the process of comparing people to the process of comparing stars; and introduced various charateristics upon which the comparison of stars is based.

### *Outcome measure*

The outcome measure, termed the astrology test, consisted of twenty-four multiple choice items. In developing the astrology test the following procedures were undertaken. First, a table of specifications was constructed which took into account the content of the learning passage and used knowledge and comprehension as broad objectives (Bloom, 1956). Second, an item pool consisting of thirty-five items was developed and submitted to three judges to determine how well the test reflected its intended objectives. A particular item was dropped if the three judges were not in unanimous agreement as to its classification in either the knowledge or comprehension categories. This procedure resulted in a reduced pool of twenty-nine items. Third, the learning passage and outcome measure were administered to a sample of ninth grade students. Following the computation of item difficulty indices, item discrimination indices, and item correlations with the total test, five additional items were dropped. Based upon this tryout, the test yielded an internal consistency coefficient of .78.

An additional precaution was taken to guard against the possibility that information included in the organizer treatments would be directly relevant to the astrology test. The three experimental treatments (i.e. graphic organizer, proze organizer, and control) were administered to a sample of students without the learning passage. On the subsequent administration of the astrology test, the three groups did not score beyond what might have been expected on a chance basis. Therefore, it was assumed that the organizer treatments contained no information that would directly assist subjects in responding to the astrology test.

*Treatment implementation*

Administration of treatments occurred during each subject's regularly scheduled English class. Subjects were randomly assigned, within classes, to the three experimental groups. On the day of treatment implementation for their class, the subjects in each group reported to one of three special rooms. At this time, they received one of the three treatments and read the learning passage.

A maximum time of five minutes was alloted for presentation of the organizers. The subjects were allowed twenty-five minutes to read and review the learning passage. On the day following administration of treatments, students returned to their regularly scheduled English class where they responded to the astrology test.

## RESULTS

The means and variances for each of the experimental groups are presented by grade levels in Table 2. For the total group, the scores ranged from 3 to 23 with a mean of 9.50 and a variance of 14.78.

Table 2
Means (and Variances) across Treatments by
Grade Level: Astrology Examination

| Grade | Treatments | | |
|---|---|---|---|
| | Graphic Organizer | Prose Organizer | Control |
| Six | 7.77<br>(7.78) | 7.36<br>(11.42) | 8.60<br>(15.31) |
| Seven | 8.79<br>(15.96) | 8.77<br>(14.33) | 8.53<br>(10.33) |
| Eight | 8.47<br>(10.26) | 8.86<br>(13.36) | 9.56<br>(14.45) |
| Nine | 8.69<br>(8.03) | 9.57<br>(15.85) | 9.38<br>(13.66) |
| Ten | 10.35<br>(16.98) | 9.67<br>(13.63) | 9.92<br>(16.58) |
| Eleven | 12.14<br>(14.14) | 11.36<br>(14.99) | 10.55<br>(13.09) |
| Twelve | 11.31<br>(13.30) | 12.80<br>(19.31) | 11.21<br>(15.52) |

The astrology exam was given 24 hours after the readings. At each grade level, two questions were asked: is there any significant difference between the combined experimental treatments and the control group; and is there any significant difference between the two treatments. No significant differences occurred at any of the grade levels for either question.

A major limitation of the study involved recomputation of test reliability for the total sample. The K-R #20 yielded an internal consistency coefficient of .69.

## CONCLUSIONS AND RECOMMENDATIONS

The present investigation addressed the question: Are Ausubel's theoretical assertions regarding the facilitative effects of advance organizers generally applicable? Insofar as learning tasks of the type implemented in this study are concerned, the answer appears to be no. Apparently, within-grade student variability precludes large scale usage of advance organizers as instructional tools.

Subsequent studies, in keeping with the previously described strong inference process, should focus upon the question: Who, if anyone, are advance organizers appropriate for? In the opinion of the present investigator, such studies should be restructed to consideration of personological variables which are readily identifiable by classroom teachers and useful as a means of differentiating instruction.

## REFERENCES

Ausubel, D.P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal Educational* Psychology, 1960, *51*, 267-272.

Ausubel, D.P. *The psychology of meaningful verbal learning*. New York: Grune and Stratton, 1963.

Ausubel, D.P. *Educational psychology: A cognitive view*. New York: Holt, Rinehart, and Winston, 1968.

Ausubel, D.P. Personal correspondence to Barron, Sept. 30, 1969.

Ausubel, D. P., & Fitzgerald, D. Organizer, general background and antecedent learning variables in sequential verbal learning. *Journal of Educational Psychology*, 1962, *53*, 243-249.

Ausubel, D.P., & Youssef, M. Role of discriminability in meaningful parallel learning. *Journal of Educational Psychology*, 1963, *54*, 331-336.

Barron, R.F. The use of vacabulary as an advance organizer. In H.L. Herber and P.L. Sanders (Eds.), *Research on reading in the content area: First year report*. Syracuse, N.Y.: Syracuse University Press, 1969.

Baumman, D.J., & Glass, G.V. The effects on learning of the position of an organizer. Paper presented at the meeting of the American Educational Research Association, February, 1969.

Bloom, B.S. (Ed.), *Taxonomy of educational objectives: Cognitive domain*. New York: David McKay Company, 1956.

Campbell, D.T., & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963.

Carter, C.E., Proger, B.B., Mann, L., & Haughey, C.E. The relationships between advance organizers for sixth-grade reading selections, reading level, and sex. Paper presented at the meeting of the American Educational Research Association, Minneapolis, February 1970.

Deitz, D. *Stars and the universe*, New York: Random House, 1968.

Estes, T.H., Mills, D.N., & Barron, R.F. Three methods of introducing students to a reading-learning task in two content subjects. In H.L. Herber and P.L. Sanders (Eds.), *Research on reading in the content area: First year report*. Syracuse, N.Y.: Syracuse University Press, 1969.

Fitzgerald, D., & Ausubel, D.P. Cognitive versus affective factors in the learning and retention of controversial material. *Journal of Educational Psychology*, 1963, *54*, 73-84.

Grotelueshen, A., & Sjogren, D.D. Effects of differentially structured introductory materials and learning tasks on learning and transfer. *American Educational Research Journal*, 1968, *5*, 191-202.

Jerrolds, B.W. The effects of advance organizers in reading for the retention of specific facts. (Doctoral dissertation, University of Wisconsin) Dissertation Abstracts, May 1967, *28*, 4532A. Ann Arbor: University Microfilms, No. 67-16,963.

Murphy, H.D. The effect of different locations of chapter summary on immediate and delayed recall. (Doctoral dissertation, George Peabody College for Teachers) Dissertation Abstracts, 1962, *23*, 2014. Ann Arbor: University Microfilms, No. 62-5682.

Platt, J.R. Strong inference. *Science*, 1964, *146*, 347-353.

Proger, B.B., Mann, L., Taylor, R.G., Morris, V.R., & Reckless, D.E. Advance and concurrent conceptual structuring mechanisms in detailed reading passages. Paper presented at the meeting of the American Educational Research Association, Minneapolis, February 1970.

Proger, B.B., Taylor, R.G., Mann, L., Coalson, J.M., & Bayuk R.J. Conceptual prestructuring for detailed verbal passages. *Journal of Educational Research*, 1970, *64*, 28-34.

Scandura, J.M., & Wells, J.N. Advance organizer in learning abstract mathematics. *American Educational Research Journal*, 1967, *4*, 295-301.

Smith, R.J., & Hess, K.D. The effects of prereading assistance on the comprehension and attitudes of good and poor readers. *Research in the Teaching of English*, 1969, *3*, 166-177.

Thelen, J.N. The use of advance organizers and guide material in viewing science motion pictures in the ninth grade. (Doctoral dissertation, Syracuse University) Dissertation Abstracts, June 1971, *31*, 6288A. Ann Arbor: University Microfilms, No. 71-10,990.

Triezenberg, H.J. Testing for concepts of ecological systems. Wisconsin Research and and Development Center for Cognitive Learning, Practical Paper No. 3. Madison: University of Wisconsin, 1968.

THOMAS H. ESTES
University of Virginia

# EFFECTS OF ADVANCE ORGANIZERS UPON MEANINGFUL RECEPTION LEARNING AND RETENTION OF SOCIAL STUDIES CONTENT[1]

In his recent book, *Undestanding Reading*, Frank Smith (1971) states that what the brain tells the eye is accountable for much more of comprehension than what the eye tells the brain. That is, it is not so much the nature of what is to be read (what is seen by the eye) as it is the nature of the reader (the information processing activity of which he is capable) which determines comprehension. One implication of this suggests the possibility that comprehension may be more affected through modification of the reader in preparation for reading than through modification of the reading material.

Similar thinking must have led Ausubel (1960) to formulate a theory of comprehension which he operationalized as an advance organizer. This theory is comprised of three successively dependent hypotheses:

1) A person's cognitive structure is an intricate system of concepts, hierarchiacally arranged in terms of their inclusiveness.

2) New concepts and understandings are learned insofar as they can be subsumed into the learner's hierarchical congnitive structure.

3) Learning and retention are facilitated by a conscious and active awareness of the proper subsuming concepts within which the new learning fits.

Ausubel's operationalization of this theory involves preceding the learning task (in most cases a printed passage to be read and understood) with a short introductory passage dealing with the content of the learning task at a higher level of generality and inclusiveness. It is Ausubel's idea that this introductory activity will serve to mobilize relevant concepts in the mind of the reader under which the content of the passage can be subsumed. The theory and its operationalization are logically sound and stand on a rather firm empirical base (Ausubel, 1960; Ausubel and Fitzgerald, 1961; 1962; Ausubel and Youssef, 1963; Earle, 1969; Estes, Mills, and Barron 1969; Fitzgerald and Ausubel, 1963 Grotelueshen and Sjegren, 1968; Murphy, 1962).

---

[1] A note of thanks is due to the staff and students of the Green County Public Schools, Stanardsville, Virginia, who participated in the study, and to the researcher's assistant, Miss Julie Johnstone, whose help made this study possible.

Several difficulties with the advance organizer have become apparent, however. As an aid to comprehension, it is almost impossible to use. The teacher or researcher is never quite sure whether the introductory passage is at a truly higher level of generality and inclusiveness in comparison to the learning passage. Furthermore, one can never conveniently know the nature of the concepts the organizer is supposed to mobilize in the mind of the reader, or, indeed, whether the concepts even exist for the individual. In fact, organizers seem definable only on an *ex post facto* basis. If it worked, it was an advance organizer for the reader; if it didn't, it was not. Ausubel himself has stated this (1968). It seems that if the learner's cognitive structure is incomplete in terms of the necessary relevant concepts, there will simply be nothing for the organizer to organize. On the other hand, if the learner's understanding of those concepts is already very clear, the organizer will act as mere noise, either having no effect or actually inhibiting learning which might otherwise have been successful.

In response to the two problems listed above, Barron (1969) has developed a prereading technique which is theoretically similar to the advance organizer but which allows for an interaction between the learner and teacher. The technique is described as a structured overview. It is constructed by arranging words relevant to the important concepts in the learning passage in a graphic form to depict for the reader the relationships between those concepts and the general area of knowledge of which the understandings in the passage are a subpart. The interaction between the learner and teacher allows the latter to estimate the relevancy of the concepts to the learner's existing cognitive structure and to make minor adjustments in the strutcured overview as it is presented. Hopefully, the teacher can help the learner appreciate this relevancy while simultaneously aiding him in the mobilization of the concepts which he will find useful in understanding the learning task.

Part of the second problem listed above is not completely solved by the structured overview, however. A question remains as to whether the learner's cognitive structure is either incomplete or already very stable and clear with respect to the concepts necessary for subsumption of undrstandings contained in the learning passage. It is possible that the success of any prereading organizing device depends on the relationship between the difficulty of the learning passage and the reading ability of the learner.

This idea is based on two unproven but related and logically apealing ideas. First, reading ability is probably closely related to the reader's previous conceptualizations. Second, it is those conceptualizations which, when properly mobilized and available, allow the reader to comprehend whatever he reads. The implication of this for the success of prereading organizing devices is that where the difficulty of a passage is within the range of ability of the reader, these devices will act as facilitators of increased comprehension. On the other hand, where the difficulty is outside the range of ability of the reader (whether too difficult or too easy), prereading organizers will not have this facilitating effect.

This study sought, therefore, to answer two questions: does the structured overview as proposed by Barron have the same or similar effects on reading comprehension as Ausubel's advance organizer?; is the facilitative effect of both devices a function of the relationship between the difficulty of the learning passage and the reading ability of the learner?

## METHOD

Subjects for this study were drawn from a small, rural community high school. Two social studies classes at each grade level, 8-12, were involved in the experiment. All students were asked to read a 3,000 word passage on the topic of the rise of labor unions in the United States. The grade level difficulty of the passage was 9.6 as measured by the Dale-Chall readability formula. The particular topic of the passage was judged to be rather unfamiliar to the students since their rural culture does not include concern with labor unions.

Students (N=157) were randomly assigned within classes to three treatment conditions. One group of students was asked to read the labor union passage preceded by a shorter passage concerning the more general topic of industrialization. This was the advance organizer. It, too, was written at a ninth grade level of difficulty. It described industrialization as possible because of changes in transportation, improved machinery, increased efficiency in production, the rise of factories, growing corporations, and increased numbers of workers. The topic of labor unions can be thought of as a subpart of the more general topic, industrialization.

A second group of students were asked to read the same passage preceded by a short discussion of industrialization stimulated by structured overview. The content of the overview was essentially the same as that of the advance organizer, the difference being that the concept and its accompanying terminology were presented in graphic form. Students were encouraged to discuss the overview by relating the terms in it to their own store of experience and knowledge. The discussion was limited to approximately ten minutes.

The third group of students who read the passage served as a control group. No prereading activity was provided for them.

A test over the content of the learning passage was administered to all pupils twenty-four hours after treatment. A forty-item, first dratf of the test had been administered to three groups of ninth grade pupils who were different from those later to serve in the experiment. One group took the test after having read the learning passage, a second group took it after having had exposure to only the advance organizer, and a third group took it after having only discussed the structured overview. Items retained in the final draft were those which were easier for the group who read the passage than for either the group who read the organizer or those who discussed the structured overview. This was necessary in order that the content of the passage and not the content of either the organizer or overview be reflected in the criterion instrumnt. Twenty-two items were retained for the final draft which had a split-half reliability of +.75.

In addition to the data collected in this experiment, previously collected information was assimilated on the pupils' reading ability. An all-school testing program, administered in March of the pupils' eighth grade year, served as a source for this data. Very few pupils were lost from the experiment due to missing data since the population of the school tends to be very stable. Though interpolations to their present reading level from information which was three months to four years old could not be assumed accurate in terms of assigning a grade equivalency to an individual, it was felt that relative standing of the pupils would have remained fairly constant. For example, high ability readers in eighth grade probably remain high relative to their peers at a later date. Since what was needed for this experiment was an estimate of relative standing for each pupil, the data seemed sufficient.

Based on this information, subjects were divided into four reading level groups. Interpolations suggested that Group 1 was functioning at below a grade 7.0 reading ability level, Group 2 at 7.0 - 8.9 level, Group 3 at a 9.0 - 10.9 level, and Group 4 at a level above 10.9.

## RESULTS

The empirical outcome of this study is depicted in Figure 1. A mean score on the criterion instrument (the test over the labor union passage) was computed for each experimental group for each of the four reading level groups. No statistically significant differences appeared among the treatment conditions at any reading ability level. The plot does reveal, however, that though differences are small, a pattern to the scores is evident. The structured overview treatment group performed consistently better than the advance organizer group at each ability level. This consistency of results across ability levels suggests an affirmative answer to the first question posed earlier—the structured overview and the advance organizer seem to have had similar effects on reading comprehension. Performance of the control group, on the other hand, was opposite to what was expected. Their scores were lower or equal to the experimental groups' when reading ability was high or low, but where reading ability was most commensurate with the difficulty of the reading passage, control scores were higher than those of either treatment group.

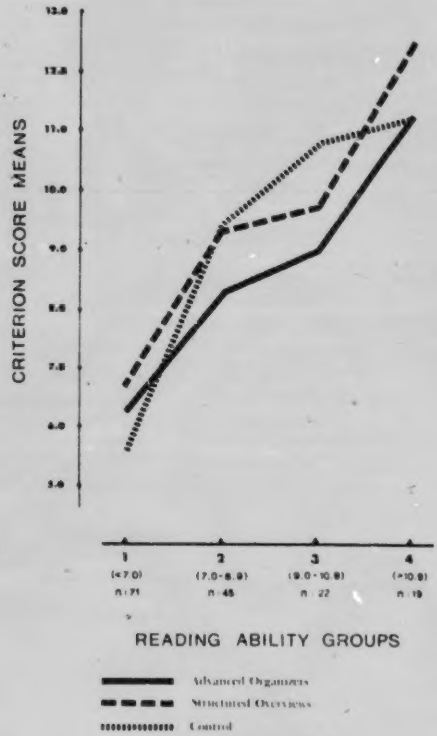


*Fig. 1. Criterion scores plotted against reading ability groups.*

## DISCUSSION

The absence of statistically significant results in an experiment is always difficult if not impossible to interpret. Final judgment concerning relative effectiveness of treatments must, of course, be withheld. But results demand exploration if only in the form of conjecture and in formulation of further plans to subject the theory to empirical test. Therefore, while this study failed to offer strong evidence in favor of its theoretical base, it does provoke several thoughts.

At least four factors may have mitigated against the success of this study in achieving statistically significant results. First, the criterion instrument was quite difficult, this tended to keep all mean criterion scores low. Second, the reading ability of pupils involved was not distributed evenly across their grade levels—almost three-quarters of the sample was reading below an approximate ninth grade level of ability as interpolated from earlier testing. Perhaps data on reading ability collected at a time closer to that of the experiment would have revealed a different pattern, though this is questionable. Third, and related to this, the small number of pupils in the top reading ability groups worked against statistically meanigful findings. Fourth, and perhaps most important, it is possible that rough groupings of pupils at four approximate ability levels, while compensating for large variance of reading levels within grade levels, may have introduced another source of crucial variance in the form of chronological age. Age, like reading ability, is probably also closely related to cognitive structure and informational background.

A subsequent study should be designed to compensate for this variance by introducing chronological age as a third independent variable. The hypothesis posited in this case would be that where age (or grade in school) and reading ability were both in accordance with the informational content and difficulty of the reading passage, an organizing device would operate in one way, but when three-way match was out of kilter, it would operate in some other pattern.

What might be that pattern? Will the structured overview continue to function at a more facilitative level than the advance organizer as in the present study? Will the devices continue to have similar effects? Will the control group, with no prereading organization, again function better at mid-range ability (and grade) levels but not at the two extremes? The next step in this line of investigation will seek answers to these questions.

## REFERENCES

Ausubel, D. P. The use of advance organizers in the learning and retention of meanigful verbal material. *Journal of Educational Psychology*, 1960, *51*, 267-272.

Ausubel, D. P., & Fitzgerald, D. The role of discriminability in meaningful verbal learning and retention. *Journal of Educational Psychology*, 1961, *52*, 266-274.

Ausubel, D.P., & Fitzgerald, D. Organizer, general background, and antecedent learning variables in sequential verbal learning. *Journal of Educational Psychology*, 1962, *53*, 243-249.

Ausubel, D. P. Telelecture to Syracuse University, Nov. 1968.

Ausubel, D. P., & Youssef, M. Role of discriminability in meaningful parallel learning. *Journal of Educational Psychology* 1963, *54*, 331-336.

Barron, R. F. The use of vocabulary as an advance organizer. In H. L. Herber and P. L. Sanders (Eds.), *Research on Reading in the content area: First year report.* Syracuse, N.Y.: Syracuse University Press, 1969.

Earle, R. A. Use of the structured overview in mathematics classes, In H. L. Herber and P. L. Sanders (Eds.), *Research on Reading in the content area: First year*

Estes, T. H., Mills, D. N., & Barron, R. F. Three methods of introducing students to a reading-learning task in two content subjects. In H. L. Herber and P. L. Sanders (Eds.), *Research on reading in the content area: First year report.* Syracuse, N.Y.: Syrause University Press, 1969.

Fitzgerald, D., & Ausubel, D. P. Cognitive versus affective factors in the learning and retention of controversial material. *Journal of Educational Psychology,* 1963, *54* 73-84.

Grotelueshen, A., & Sjogren, D. D. Effects of differentially structured materials and learning tasks on learning and transfer. *American Educational Research Journal,* 1968, *5*, 191-202.

Murphy, H. D. *The effect of differet locations of chapter summary on immediate and delayed recall.* (Doctoral dissertation, George Peabody College for Teachers.) Dissertation Abstracts, 1962, *23*, 2014. Ann Arbor, Michigan: University Microfilms, 1962. No. 62-5682.

Smith, F. *Understanding reading.* New York: Holt, Rinehart, and Winston, 1971.

BOB W. JERROLDS
University of Georgia

# EFFECTS OF ADVANCED ORGANIZERS IN READING FOR THE RETENTION OF SPECIFIC FACTS

There are many purposes for reading, one of the most significant being to remember specific facts. Many, if not most, classroom tasks require the student to identify spicific facts from reading assignments and to retain those facts for later use. Although he may be told to read to remember those facts, the student is not taught how this task can be accomplished. Indeed, although teachers often make such assignments, little is actually known about the process itself or the most effective techniques for its accomplishment.

Ausubel (1960) has advanced the hypothesis that the most effective way of increasing learning and retention of meaningful verbal materials is by manipulating the organization, stability and clarity of the concepts relating to new material that are already established in the learner's cognitive structure. Ausubel suggests that this cognitive structure manipulation can be best accomplished by the use of cognitive organizers. He has formulated the concept of the Advance Organizer (AO) which is a set of related material, presented prior to the new material and written on a higher level of abstraction, inclusiveness and generality. Expository material can be written to contain an explicit statement of the main idea of the passage. These comprehensive statements would have many similarities to the AOs as they are conceived by Ausubel. The writing of AOs in the manner that Ausubel suggests would necessitate massive revisions in school materials. The present writer, therefore, wanted to determine whether a modified advance organizer (MAO), a single explicite statement of the main idea of a passage, could serve as effectively as the AO if students were instructed in the use of these types of advance organizers. The present study, which involved ninth grade subjects of average and above average intelligence level, examined the effects upon retention of specific facts of both the AO and the MAO under conditions of instruction and no instruction.

With regard to instruction, this writer attempted to teach certain of the subjects to use some of the techniques Ausubel says will facilitate meaningful verbal learning. It was not postulated that the ninth grade student could or should master Ausubel's theoretical construct of meaningful verbal learning. However, this writer felt that many of the processes and factors Ausubel considers essential for effective learning could be taught to ninth grade students in one session. It was held that the student could profit more from the use of AOs and MAOs if he knew more about how he was to use them.

The study had three major purposes: to determine whether advantages of the AO that have been demonstrated by Ausubel with college age subjects are similarly advantageous for ninth grade students; to examine the effects of a short program of instruction based upon the principles advocated by Ausubel for maximizing meaningful reception learning, under conditions using the AO and the MAO; and to determine the differential effects of AOs and MAOs, under conditions of instruction and no instruction, on two IQ groups, average and above average.

## METHOD

### *Subjects*

Subjects were selected from a population of 300 ninth grade students enrolled in a mid-west suburban high school. The school is typical of the suburban schools of major American cities. The average student of this junior high school is somewhat above the national norms in intelligence and socio-economic background.

The present writer sought to determine the effects of written advance organizers when subjects read a passage of task material. Obviously if the subject could not read, his performance on the test would confound the results of the study. To eliminate this factor, all students reading below the 40th percentile on the reading comprehension and vocabulary sections on the Iowa Test of Basic Skills were dropped from the sample population. From among the ninth graders who scored at or above the 40th percentile on the Iowa subtests, a pool of 83 students who schored 100 to 110 and a pool of 61 students who scored 115 and up on the California Test of Mental Maturity were identified. Seven students from each pool were assigned at random to each of the conditions of the experiment. One student assigned to each cell was an extra subject added to guard against attrition due to absence. Thus the total number of students participating in the experiment was 98. Extra subjects remaining after attrition were randomly dropped, leaving 84 subjects, six in each cell.

*Procedures, Materials and Treatment Groups*

Table 1 shows the treatments and chronology of the experiment:

Table 1

Overview of Treatments and Chronology

| Group | Wednesday | Thursday | Friday | | Monday |
|---|---|---|---|---|---|
| | Type of Instruction | Type of Subsumer | Task Material | | Test |
| 1 | Instruction in the use of the MAO as a subsumer | MAO | Task Material | Three day interval | Test of Delayed Retention |
| 2 | No Instruction | MAO | | | |
| 3 | Instruction in the use of the AO as a subsumer | AO | | | |
| 4 | No Instruction | AO | | | |
| 5 | No Instruction | Buffer Material | | | |
| 6 | No Instruction | AO | No Task Material | | |
| 7 | No Instruction | No Subsumer | No Task Material | | |

Only two groups were given instruction: Group 1 was instructed in the use of the MAO as a subsumer and Group 3 was instructed in the use of the AO as a subsumer. Each group received 55 minutes of instruction. The processes they were to use in maximizing the learning and retention of specific facts were explained to each group. The processes, all derived from the basic Ausubel principles, were presented to the subjects in as simple a way as possible.

All materials used in this study were carefully constructed to meet the needs of the study and to conform to eighth and ninth grade reading levels. Materials were pilot tested in another suburban junior high school. No particular difficulties in vocabulary or other elements of readability were revealed. A 1,000-word passage on life in the ocean was used to demonstrate the techniques and procedures of employing a subsumer. Both groups were given the same passage in this instruction period. Instruction was very similar for the two groups with the only variations being those dictated by the nature of the two types of subsumers. No test was given on the material used in the instruction.

On the day following the instruction, Groups 1 and 2 were given MAOs for a 1,000-word passage on espionage organizations. Groups 3, 4 and 6 were given an AO designed for that same passage. Group 5 was given the buffer material, a passage of equal length to the AO on the history of espionage. This buffer material was written on the same level of abstraction, generality and in clusiveness as the task material. It contained no facts or generalizations that would be helpful to subjects in dealing with the task material on the test of delayed retention. The purpose of buffer passage and procedure was to obviate the problem of learning set in relation to the control group and the experimental groups. Group 5 received no subsumers. Groups 6 and 7 were included in this study to insure that the AO contained no specific information that would aid in answering the questions on the test of delayed retention. If it is true that the AO contains no such specific information, Group 6 receiving only the AO and the test of delayed retention should not be significantly different from Group 7 which received only the test of delayed retention.

On the day following the presentation of the AOs, MAOs, and buffer material, all groups in the experiment except Groups 6 and 7 were presented with the 1.000-word passage on espionage organizations. Subjects were given fifteen minutes in which to study the material in order to remember the specific facts it contains. They were told not to rush since the time was to be sufficient for the task.

After a three-day interval, the test of delayed retention was administered to each group in the experiment. The purpose of this study was to examine the retention of specific facts; therefore, test items were constructed around the specific facts in the task material. Each answer would complete a statement, thus constituting one specific fact from the task material. All distractors were plausible and were of the same general grammatical structure and length as the answer. There were four possible answers for each test item. To avoid inadvertent patterning, these answers and distractors were randomly assigned to the 1st, 2nd, 3rd, and 4th positions.

Validity of the test of delayed retention was maintained by defining in advance the exact nature of content to be sampled, then constructing the test items in accordance with the specifications. Three ninth grade teachers were asked to judge as to whether the test did indeed measure the retention of specific facts as defined in this study. Three other teachers, uninformed as to the nature of the test, were given the task material and test and asked to decide what was being tested with this particular instrument and to determine whether the items would require approximately the same level of understanding. In answer to the question about the general nature of the test, these teachers concluded that the test measured only factual information. Specific items held by these teachers to be decidedly above the other items in the level of understanding required were eliminated from the test or reconstructed. The test was piloted; some items on the test of delayed retention were eliminated and others reconstructed on the basis of this piloting since some items showed a low discrimination potency and certain proposed distractors did not seem to function adequately. On the final test, the use of K-R formula revealed a reliability of .68.

### *Design and Analysis*

Twelve subjects were requisite for each of the seven treatments, six within the IQ range of 100-110 and six within the range of 115 up. To guard against attrition, one extra subject was assigned to each cell. After attrition the remaining extra subjects were randomly dropped. The subjects' performance was measured by the twelve-item multiple choice test of delayed retention.

Not all possible comparisons among the seven treatment groups were made. Only the specific comparisons suggested by the theoretical bases of the experiment were examined. Hays suggests that the technique of planned comparisons be used instead of the ordinary analysis of variance and F test when ". . . instead of planning to analyze his data to see if any over-all experimental effects exist, the researcher at the out-ste has a number of particular questions he wants to answer separately [Hays, 1963, p. 460]." For analysis, planned comparisons are simply individual comparisons made by means of the *t* statistic. The groups thus compared in this study were: 3 x 4, 1 x 2, 3 x 1, 1 x 4 and 6 x 7. The .05 level of significance was accepted in all comparisons.

The experimenter wanted to consider other comparisons. However, the *t* statistic could not be used further without danger of using the data on one group so many times as to bias the estimate of the experimental error. Group 5 was a control group. All comparisons made against the control group were made using the Dunnett *t* statistic. This statisitc sets a higer critical value for significance than the *t* statistic and is a special method for handling multiple comparisons against one control group. Winer says of this test, "The approach used by Dunnett is a special case of the problem of handling multiple comparisons by constructing a joint confidence interval on the set of all relevant comparisons [1962, p. 89]."

## RESULTS OF THE STUDY

### *Planned Comparisons*

None of the comparison groups differed at the level of significance set for this study. Since the group with instruction did not perform significantly better than the groups given no instruction, the general hypothesis of the study, that the presentation of AOs and MAOs to ninth grade students under conditions of instruction would result in their greater retention of specific facts than under conditions of no instruction, was not confirmed.

There was no significant difference when the group instructed in and using MAOs was compared with the group using the AO with prior instruction. Thus the AO whether under conditions of instruction or no instruction was not found to be superior to the MAO when subjects were trained in its use.

There was no significant difference between Group 6, the group not given the task material but only the AO and the test of delayed retention nad Group 7, the group given only the test of delayed retention. This finding substantiated the experimenter's contention that the AO itself contained no specific information that would aid in answering the items on the test of delayed retention.

*Multiple Comparisons Against the Control*

Groups using an MAO under conditions of instruction and no instruction and groups using AOs under conditions of instruction and no instruction did not differ significatly from the cotrol group which had the buffer material instead of an advance organizer. In every case, however, the trend favored the groups using advance organizers.

One general hypothesis of this study was that ninth grade subjects using advance organizers would retain specific facts better than ninth grade subjects not using advance organizers. Comparisons of Groups 1, 2, 3, and 4 to Group 5, the control group, failed to substantiate this hypothesis.

*The Effect of Two IQ Levels Within the Planned Comparisons*

Above average IQ subjects instructed in and using an MAO performed significantly better on the test of delayed retention than above average IQ subject using an MAO without prior instruction. Above average IQ subjects instructed in and using the MAO did not differ significantly from above average IQ subjects instructed in and using an AO nor from those using the AO without prior instruction. Above average subjects instructed in and using the AO did not differ significantly from those using the AO without prior instruction. There was no significant difference between above average IQ subjects given only the AO and the test of delayed retention and above avrage IQ subjcts given only the test of delayed retention. This finding indicated that the above average IQ subjects found no specific information in the AO to aid them in answering the items on the test of delayed retention.

With average IQ subjects none of the comparisons reached the .05 level of significance. Neither the type of subsumer, AO or MAO, nor the condition of instruction or no instruction was found to be a significantly varying factor. No significant difference was revealed between average IQ subjects given only the AO and test of delayed retention and those average IQ subjects given only the test of delayed retention.

### *The Effects of Two IQ Levels Within the Multiple Comparisons Against the Control*

No significant difference was found when above average IQ subjects of the four experimental groups were compared with above average IQ subjects in the control group. Groups using an MAO under conditions of instruction and no instruction and groups using an AO under conditions of instruction and no instruction did not differ significantly from the control groups which had buffer material instead of an advance organizer. With the exception of the group using the MAO without prior instruction, the existing differences favored the groups using advance organizers. The group instructed in and using the AO and the group using the AO without prior instruction approached significance when compared with the control group.

No significant difference was found when average IQ subject in the four experimental groups were compared to average IQ subjects in the control group. Groups using an MAO under conditions of instruction and no instruction and groups using an AO under conditions of instruction and no instruction did not differ significantly from the control group which had buffer material instead of an advance organizer. In every case the existing differences favored the groups using advance organizers, with the groups instructed in and using the AO approaching significant level.

The above average group did significantly better under conditions of instruction in the use of the MAO than the above average group using the MAO without prior instruction. In all other comparisons examining this hypothesis, the effectiveness of the AOs and MAOs did not vary within the two levels of intelligence whether under conditions of instruction or no instruction.

### REFERENCES

Ausubel, D. P. The use of advanced organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1960, *51*, 267-272.

Hays, W. L. *Statistics for psychologists*. New York: Holt, Rinehart and Winston, 1963.

Winer, B. J. *Statistical principals in experimental design*. New York: McGraw-Hill, 1962.

CLESSEN J. MARTIN
MARY ANNE HERNDON
Texas A&M University

# DEVELOPMENT OF TELEGRAPHIC PROSE BASED ON A RANDOM WORD DELETION SCHEME[1]

A computer-based random word deletion scheme was used for developing telegraphic reading materials. The possibility of developing telegraphic materials was based upon the assumption that written and spoken language contains many words and word sequences which are unnecessary for the comprehension of a message. That is, in the usual message, there are words, phrases, and sometimes even sentences which add no further information.

The fact that redundancy is present in language and that *Ss* are aware of this redundancy has been demonstrated in a number of ways. One way in which redundancy has been demonstrated was by requiring *Ss* to replace missing words in sentences. Morrison and Black (1957) found that *Ss* had little difficulty in finding words, which were appropriate substitutions for the deleted words. Garner (1962) has presented an excellent review of the language redundancy studies.

Martin and Alonso (1967) reported the results of a two-year U.S. Office of Education grant designed to examine the ability of blind children to comprehend braille prose material presented in two telegraphic styles. In this investigation, a total of 210 braille students, grades six through nine, were assigned to one of three prose treatment conditions: traditional-type prose material; medium telegraphic prose material; short telegraphic. The format of the short version was similar to that of a telegram and represented a 75% word reduction in comparison to the traditional version.

---

[1] The project reported here was performed pursuant to a contract with the U.S. Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein, however, do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

Subjects reading the telegraphic versions required significantly less time than subjects reading the traditional version. Analysis of the comprehension data revealed few significant differences among the three groups. While there was a significant decrease in reading speed on the telegraphic materials, the overall information input rate was greater for the telegraphic prose materials.

The ultimate application of the telegraphic concept to existing prose materials is dependent upon the development of objective reduction rules whereby extraneous information may be systematically eliminated from traditional materials. The present study is the first of a series of studies presently being supported by a grant from the Basic Research Program of the U.S. Office of Education. The general research strategy of the present project involves the establishment of base levels of prose comprehension when the reduction schemes are based upon random word deletion programs. Once these base levels have been established, then more logical reduction schemes will be developed in order to determine the extent to which the percentage of deleted material can be increased above the random base levels without significant decrements in comprehension.

This study was concerned with determining the effects of 10 and 30 percent random word deletion levels upon comprehension as assessed by a specially designed set relations test. In addition, comparisons were also made concerning the effects of the experimental passages upon reading rate and reading time.

## METHOD

A fictional story concerning 2 warring African nations developed by Martin and Alonso (1967) written in full prose, traditional style and containing 1620 words was used for the test material. In order to operationally define the important information in the passage, the story was written so that the central ideas of the passage were analyzable in terms of basic set relations. Important information was defined as material which described the set relations. Unimportant information was defined as narrative or descriptive material which was unrelated to the set relations. Next, two versions of the same story in which 10% and 30% of the words were randomly deleted were presented in a style similar to that of the traditional story. The deletion was done on an IBM 360-65 computer in WATFIV Fortran, using a scheme of random number generation. The actual experimental materials used by *Ss* were computer printouts. The 66 *Ss* tested in this experiment were enrolled in an undergraduate course in Educational Psychology at Texas A&M University, and were randomly assigned to one of three treatment conditions (22 Ss/treatment). The treatment conditions were: Traditional (A): 10% Deletion (B); and 30% Deletion (C). The *Ss* were instructed to read their passage at a rate which was comfortable, but not to reread any sections of the passages. Additionally, the *Ss* were told by *E* to record the time when they finished reading the selections. A large clock was provided for the *Ss* to use in recording their time.

After the *Ss* completed the reading of the passages, they were instructed to open their test booklets and answer the set relation questions. The test consisted of 20 items, 10 of which were concerned with nested relationships and 10 items with disjunctive relationships. Each item required the *S* to circle the true statement from a pair of statements, "all of the expert warriors belonged to the Pro-Lester Union," or "only some of the expert warriors belonged to the Pro-Lester Union."

A total of four dependent variables were analyzed: total number correct, reading rate, reading time, and S-score. The S-scores were based on Dawes' (1964) model, in which he postulates two types of errors possible in the recall of set relations. Pseudo-discrimination errors are made when the *Ss* make a discrimination when there is none, and overgeneralization errors are made when a *S* fails to recall that some, rather than all or no members of a set were included in another set. Dawes predicts that *Ss* will make more overgeneralization errors than pseudo-discrimination errors. The S-score was based upon the number of nested items correct minus the number of disjunctive items correct plus 10. A constant of 10 was added in order to avoid negative values. No time limit was imposed on either the reading or the testing sessions. The *Ss* recorded their responses on IBM answer sheets.

The following are excerpts from the three versions.

| Traditional Version A | 10% Deletion Version B | 30% Deletion Version C |
|---|---|---|
| In the year 1800 on the continent of Africa, two unfriendly nations existed, the nation of Mambo and the nation of Yam. The Nile River separated these two nations. | In the year 1800 on the of Africa, two unfriendly nations existed, the nation of Mambo and the nation of Yam. The river separated these two nations. | In the year 1800 on continent Africa, two unfriendly nations existed, the nation Mambo and the nation of Yam. The Nile River separated these two. |

## RESULTS

Table 1 presents the means and standard deviations of the four four dependent variables for each of the three passages.

Four 1x3 analyses of variance were performed on the data. The analysis of the total number correct on the set relations test yielded a significant *F* value of 7.81, ($df = 2/63$, $p < .05$). The Scheffé test (Ostle, 1963) was used in order to make individual comparisons on the three

Table 1
Means and Standard Deviations for the Four Dependent Variables

| Dependent Variables | Passages | | | | | |
|---|---|---|---|---|---|---|
| | Traditional (A) | | 10% Deletion (B) | | 30% Deletion (C) | |
| | $\overline{X}$ | S.D. | $\overline{X}$ | S.D. | $\overline{X}$ | S.D. |
| Total Correct | 15.01 | 1.72 | 13.59 | 2.48 | 12.50 | 2.04 |
| Reading Rate (wpm) | 217.10 | 57.10 | 173.30 | 39.50 | 147.20 | 36.90 |
| Reading Time (min) | 8.00 | 2.37 | 9.05 | 3.04 | 8.23 | 2.50 |
| S-Score | 7.91 | 1.23 | 7.95 | 2.44 | 7.50 | 1.99 |

treatment means at the .05 level. The only difference in treatment means was between versions A and C. No significant differences were found between versions A and B, and between versions B and C.

Analysis of the reading rates yielded a significant *F* value of 12.09, ($df = 2/57$, $p < .01$). The Scheffé test indicated a significant difference between groups reading versions A and C. No other differences were significant.

There was no significant difference in average reading time for the three treatment groups ($F = .85$, $df = 2/57$).

The analysis of variance on the S-scores revealed no significant difference among the three versions ($F = .36$, $df = 2/63$).

## CONCLUSIONS

The comprehension results demonstrated no significant differences between Version A and B and between Versions B and C. The results demonstrated that it was possible to randomly delete 10% of the words from the experimental passage without significantly decreasing comprehension of the basic set relationships described in the passage. However, there was 17% less comprehension associated with Version C compared to the comprehension demonstrated after reading Version A. Examination of reading rate revealed at 32% reduction in words per minute for Version C compared to the reading rate recorded for Version A. No significant difference was observed in reading rate between Versions A and B. Also, no differences were observed in total reading time for the three versions.

It was possible to present the major themes of the three versions in Venn diagram form. With this model, it was possible to operationally define the essential structure of the prose passages. Dawes (1964) found more overgeneralization errors than pseudo-discrimination errors in his prose passage. The explanation which he offered for these findings was one of information overload. Since overgeneralization results in a reduction in the number of categories in the passage, it reduces the amount of information stored in memory. Pseudo-discrimination errors, on the other hand, produce an increase in amount of information. Because of the limitation in human memory, Dawes predicted more over-generalization than pseudo-discrimination errors. Analysis of the S-scores in the present experiment revealed no significant difference among the three versions. As a matter of fact, there was a non-significant trend toward more pseudo-discrimination e rrors than overgeneralization ones. Thus, the results are ambiguous with respect to the relative imposition made upon memory among the three treatment versions.

Replication of the present study has recently been completed by Martin and Herndon (1971). Two modifications were made in the replication study. One change involved the retyping of the computer printouts so that hyphenation at the end of each printed line corresponded to normal hyphenation rules as opposed to random hyphenation is in the computer printout. A second change involved the administration of a 39-item multiple choice test in addition to the set relations test. The results of the replication study demonstrated no significant effects between the Traditional Version and Version B (10% Deletion) on any of the dependent variables. Significant effects were obtained between the Traditional Version and Version C (30% Deletion) on reading rate, comprehension of the set relations and scores on the multiple choice test. In all cases, inferior performance was associated with Version C. In the replication study the wpm reading rate was approximately 20% less for *S*s reading Version C in comparison to *S*s reading Version A. Thus, the typed format for Version C was easier to read than the printout for Version C in the present study.

In summary, 10% of the words in the prose selection used in this experiment were randomly deleted without appreciably affecting either comprehension as defined by performance on the set relations test or the multiple choice test. Also, when the 10% Version was read in the typed format rather than the printout format, there was no difference in reading rates. Thirty per cent random deletion produced significant decrements on all dependent variables except total reading time and S-scores.

## REFERENCES

Dawes, R. M. Cognitive Distortion. *Psychological Reports*, 1964, *14*, 443-459.

Garner, W. R. *Uncertainty and structure as psychological concepts.* New York: Wiley, 1962.

Martin, C. J., & Alonzo, L. *Comprehension of full length and telegraphic material among blind children.* (ER-42) East Lansing: Educational Publication Services, 1967.

Martin, C. J., & Herndon, M. A. Comprehension of telegraphic prose: effect of a random word deletion scheme. Paper presented at the meeting of the Psychonomic Society, St. Louis, Nov. 1971.

Morrison, H. M., & Black, J. W. Prediction of missing words in sentences. *Journal of Speech Disorders*, 1957, *22*, 236-240.

Ostle, B. *Statistics in research.* Ames: Iowa State University Press, 1963.

FRED W. OHNMACHT
JAMES T. FLEMING
State University of New York at Albany

# CLOZE AND CLOSURE: A SECOND ANALYSIS

A previous study (Ohnmacht, Weaver, and Kohler, 1970) examined the relationship of flexibility and speed of closure with a number of cloze tasks representing structural, lexical, concrete, and abstract deletion strategies. Measures of vocabulary and associational fluency were included in the study with the hope that separable portions of cloze variation, representing communalities with verbal and perceptual abilities, could be identified through application of factor analytic techniques. The findings, based on 113 high school Ss, indicated that the cloze tasks could be differentiated into two discriminable dimensions. Further, perceptual abilities exhibited relatively low but significant correlations with all of the cloze tasks.

More recently, we have replicated the aforementioned study with a sample of 92 college students as subjects. The first part of this paper summarizes these new data and contrasts the findings with those reported earlier. The second part of the paper considers the findings from a linguistic perspective.

## METHOD

A battery of twelve sub-tests was administered to samples of 113 high school and 92 college students. The high school sample spanned grades 9 through 12 and was drawn from a semi-rural school in Northeastern Georgia. The college sample consisted of paid volunteers enrolled in undergraduate educational psychology at SUNY at Albany. The battery consisted of four cloze forms and two measures each for speed of closure, flexibility of closure, associational fluency, and vocabulary. A brief description of each test follows. Complete descriptions of all verbal and perceptual ability measures can be found in the *Kit of Reference Tests for Cognitive Factors* (French, Ekstrom & Price, 1963).

*Speed of Closure*

The speed of closure factor is defined as the ability to unify an apparently disparate perceptual field into a single percept.

> *Gestalt Completion Test* (test suggested by the Street Gestalt Completion Test). Drawings are presented which are composed of black blotches representing parts of the objects being portrayed. The subject writes down the name of the objects, being as specific about them as he can.
>
> *Concealed Words Test* (test suggested by Thurstone's Mutilated Words). Words are presented with parts of each letter missing. The subject is to write out the full word in an adjacent space.

*Flexibility of Closure*

The flexibility of closure factor is defined as the ability to keep one or more definite configurations in mind so as to make identification in spite of perceptual distractions.

> *Hidden Patterns Test* (test suggested by Thurstone's designs). Each item consists of a given geometrical pattern in some of which a single given configuration is embedded. The task is to mark each pattern in which the configuration occurs.
>
> *Hidden Figures Test* (Adaptation of the Gottschaldt Figures Test popularized by Thurstone). The task is to decide which of five geometrical figures is embedded in a complex pattern. The difficulty level of this test is high.

*Associational Fluency*

The associational fluency factor is the ability to produce words from a restricted area of meaning.

> *Controlled Associations Test* (test adapted from Thurstone's test of this name). The task is to write as many synonyms as possible (up to 12) for each of four words.
>
> *Association IV* (a test by J. P. Guilford). The task is to produce a word that is associated with both the two given words but which has different meaning in its relationships to each of them.

*Verbal Comprehension*

The verbal comprehension factor is defined as the ability to understand the English language.

> *Wide Range Vocabulary Test* (adopted from Cooperative Vocabulary Test). This is a five-choice synonym test having items ranging from very easy to very difficult.

*Advanced Vocabulary Test* (adopted from a test by J. B. Carroll.) This is a four-choice synonyms test consisting mainly of difficult items.

### *The Cloze Tests*

All cloze tests were pretests in that the Ss did not read the relevant passages in their intact form before attempting to complete the deleted units. A standard underline of 15 spaces was substituted for deleted words, and Ss recorded their responses directly on the test pamphlets.

The cloze materials were constructed from materials which had previously been ordered by difficulty level by means of the cloze procedure (Miller & Coleman, 1967). The 30, 150-word passages were divided into four groups of nine passages each. Sets of three sentences each were drawn from each of these four levels of difficulty until the required number of deletions were obtained.

*Structural cloze.* This is an "any-word" deletion which follows Taylor's (1954) original technique. Every fifth word is deleted regardless of its grammatical relationship.

*Lexical cloze.* This deletion is constructed by first identifying all the nouns, verbs, and adjectives in the passage and then eliminating every fifth noun, verb, or adjective.

*Abstract nouns.* Abstract nouns describe "objects" that cannot be perceived by the senses. Abstract nouns in the passages were identified first, and then a mechanical fifth-word deletion of these nouns carried out.

*Concrete nouns.* Concrete nouns were first identified (by choosing those nouns which did describe objects perceivable by the senses) and every fifth concrete noun deleted.

Data obtained from both samples were separately intercorrelated and subjected to a principal components solution. The original high school study reported two, three, and four factor solutions rotated to both varimax and maxplane criteria. This report contrasts the three factor solutions for each set of data after rotation of the principal component structure to the varimax criterion.

Tables 1 and 2 summarize the descriptive statistics including correlations, means, and standard deviations for the college and high school samples respectively.

Inspection of Tables 1 and 2 discloses marked differences between the systems of correlations for the high school and college samples. For the high school sample 65 of 66 correlations are significant ($p < .05$) whereas for the college sample 32 of 66 are significant. Further, of 16 perceptual ability correlations with cloze tasks, 15 are significant for the high school sample, and 5 for the college sample. Similarly, all of the verbal ability measures are significantly related to the cloze tasks for the high school sample whereas 10 of 16 are for the college sample. Perhaps more importantly, the magnitude of the relationships for the

Table 1

Correlations among variables with associated means and standard deviations (College sample, N=92.)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | X | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Cloze I-Structural | – | 55 | 33 | 36 | 17 | 15 | 28 | 22 | 21 | 25 | 38 | 47 | 23.62 | 6.44 |
| 2. Cloze II-Lexical | | – | 16 | 38 | –04 | –06 | 12 | 06 | 11 | 11 | 23 | 35 | 12.90 | 3.88 |
| 3. Cloze III-Concrete | | | – | 19 | 29 | 14 | 27 | 26 | 12 | 32 | 13 | 20 | 21.79 | 5.11 |
| 4. Cloze IV-Abstract | | | | – | 00 | –06 | 09 | –05 | 05 | 09 | 27 | 25 | 20.02 | 4.61 |
| 5. Gestalt Completion | | | | | – | 37 | 43 | 26 | –08 | 22 | –08 | –05 | 16.68 | 3.83 |
| 6. Concealed Words | | | | | | – | 40 | 20 | 03 | 20 | 06 | 04 | 23.76 | 5.88 |
| 7. Hidden Patterns | | | | | | | – | 52 | 05 | 23 | 03 | 08 | 66.11 | 16.51 |
| 8. Hidden Figures | | | | | | | | – | 10 | 10 | –08 | 03 | 14.82 | 8.77 |
| 9. Controlled Association | | | | | | | | | – | 13 | 34 | 41 | 30.42 | 8.82 |
| 10. Associations IV | | | | | | | | | | – | 30 | 31 | 9.89 | 4.25 |
| 11. Wide Range Vocabulary | | | | | | | | | | | – | 78 | 19.22 | 6.50 |
| 12. Advanced Vocabulary | | | | | | | | | | | | – | 25.68 | 5.01 |

$r > .20$, $p < .05$. $r > .26$, $p < .01$.

## Table 2

Correlations among variables with associated means and standard deviations (High school sample, N=113)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | X | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Cloze I-Structural | – | 42 | 48 | 45 | 24 | 28 | 27 | 40 | 50 | 38 | 42 | 34 | 14.56 | 7.84 |
| 2. Cloze II-Lexical | | – | 42 | 52 | 24 | 25 | 15 | 22 | 44 | 34 | 48 | 45 | 7.53 | 4.03 |
| 3. Cloze III-Concrete | | | – | 49 | 27 | 20 | 24 | 32 | 44 | 36 | 39 | 37 | 16.22 | 7.21 |
| 4. Cloze IV-Abstract | | | | – | 24 | 32 | 21 | 28 | 37 | 44 | 53 | 42 | 12.51 | 5.58 |
| 5. Gestalt Completion | | | | | – | 42 | 45 | 38 | 38 | 28 | 26 | 28 | 16.34 | 5.81 |
| 6. Concealed Words | | | | | | – | 55 | 46 | 39 | 35 | 33 | 36 | 13.80 | 4.29 |
| 7. Hidden Patterns | | | | | | | – | 47 | 20 | 18 | 32 | 22 | 42.93 | 22.06 |
| 8. Hidden Figures | | | | | | | | – | 32 | 32 | 43 | 31 | 4.93 | 5.11 |
| 9. Controlled Associations | | | | | | | | | – | 45 | 40 | 42 | 14.72 | 7.82 |
| 10. Associations IV | | | | | | | | | | – | 38 | 45 | 4.80 | 3.61 |
| 11. Wide Range Vocabulary | | | | | | | | | | | – | 69 | 7.41 | 5.75 |
| 12. Advanced Vocabulary | | | | | | | | | | | | – | 4.37 | 4.60 |

$r>.24$, $p.<.05$.

college sample are considerably less than for the high school group. Within cloze task correlations are also lower for the college group. In general the means for the college sample are substantzially higher than for the high school group without exception.

The foregoing summary is supported by the components analyses of the respective correlation matrices. For the high school sample only two latent roots exceeded unity whereas three roots exceeded unity for the college sample again indicating greater positive manifold for the high school sample. Table 3 presents three factor solutions rotated to the varimax criterion for each of the samples.

Table 3

Loadings for Three Factor Solutions
High School and College Studies Juxtaposed

| Variable | Factors | | | | | |
|---|---|---|---|---|---|---|
| | I | | II | | III | |
| | hs | c | hs | c | hs | c |
| 1. Cloze I-Structural | 75 | 69 | 22 | 31 | 14 | 07 |
| 2. Cloze II-Lexical | 56 | 81 | 02 | 00 | 50 | 17 |
| 3. Cloze III-Concrete | 75 | 30 | 14 | 50 | 17 | 10 |
| 4. Cloze IV-Abstract | 56 | 74 | 10 | -04 | 50 | 33 |
| 5. Gestalt Completion | 27 | -02 | 68 | 73 | 04 | -09 |
| 6. Concealed Words | 13 | -20 | 74 | 63 | 26 | 16 |
| 7. Hidden Patterns | 05 | 14 | 83 | 79 | 10 | 02 |
| 8. Hidden Figures | 24 | 08 | 67 | 65 | 20 | -05 |
| 9. Controlled Associations | 69 | -03 | 27 | 01 | 18 | 66 |
| 10. Associations IV | 47 | 04 | 22 | 39 | 40 | 48 |
| 11. Wide Range Vocabulary | 24 | 24 | 25 | -08 | 81 | 84 |
| 12. Advanced Vocabulary | 20 | 33 | 19 | 00 | 84 | 83 |

Inspection of Table 3 indicates that the three factor solution for the two samples bear certain similarities to each other. The first factor, in both instances, is a general cloze factor. The major discrepancy results from the altered loadings for the associational fluency tasks. For the high school samples Controlled Associations and Associations IV have major loadings on this factor whereas for the college sample they are nil. This reflects, in a parsimonious manner, the lack of dependence of cloze performance upon associational abilities for the older sample.

The second factor, in both analyses, is a perceptual factor, with all perceptual tests exhibiting their highest loadings. For the high school sample, the relationships between perceptual abilities and cloze is masked in a varimax solution. A rotation to the maxplane criterion (see table 4) produced a correlation of .47 between the two factors with the Structural and Concrete cloze tests having the highest loadings on the cloze factor. A similar finding is demonstrated in the college sample with the Concrete cloze task loading directly on the perceptual factor. The decreased positive manifold in the college group has produced a system of correlations to which an orthogonal set of factors demonstrates a superior "simple structure" fit. It should be pointed out, however, that the underlying correlations of closure tests to cloze performance are lower for the college group.

Table 4

Maxplane[1] Solution for High School Sample

| Variable | Factors | | |
|---|---|---|---|
| | I | II | III |
| 1. Cloze I-Structural | 64 | 05 | -43 |
| 2. Cloze II-Lexical | 81 | -19 | 04 |
| 3. Cloze III-Concrete | 68 | -05 | -40 |
| 4. Cloze IV-Abstract | 80 | -10 | -04 |
| 5. Gestalt Completion | 07 | 67 | -16 |
| 6. Concealed Words | 11 | 74 | 09 |
| 7. Hidden Patterns | -10 | 88 | 04 |
| 8. Hidden Figures | 18 | 64 | -03 |
| 9. Controlled Associations | 61 | 11 | -36 |
| 10. Associations IV | 63 | 05 | -05 |
| 11. Wide Range Vocabulary | 76 | 07 | 40 |
| 12. Advanced Vocabulary | 76 | 00 | 45 |

[1] Factor I and II correlate .47. Factor III has essentially zero correlation with I and II.

Factor III is a verbal ability factor dominated in both instances by the vocabulary tests. For the high school sample Lexical and Abstract Cloze tasks loaded on this factor due to the relatively strong correlations of these tasks with the verbal ability measures. The maxplane (Table 4) rotation represents these results in a different way but can be interpreted in the same manner. Factor III for the college sample represents a cleaner break of the verbal abilities from both cloze and closure abilities. Thus, for the college sample, Lexical and Abstract cloze demonstrate markedly less dependence on the verbal ability measures.

## DISCUSSION

In summary, both data sets indicate that the cloze tasks are factorially complex in that some tasks load on two factors. The high school data demonstrate a stronger positive manifold with cloze performance demonstrating greater dependence on vocabulary, associational fluency, and to a lesser extent upon closure abilities. The data seem to be consistent with a differentiation hypothesis with the data system elements displaying a general decrease in the magnitude of the intercorrelations when the college sample is contrasted with the high school group. Rankin (1965) suggested that closure factors would be more important in the performance of cloze tasks when contrasted with older Ss. These data, although trending in that direction, are not convincing. A study employing elementary school Ss, similar to those reported here, would be useful in further evaluating such a hypothesis.

Increased emphasis might be profitably placed on re-examining the assumtpions underlying the construction and use of certain cloze tests, as well as exploring a broader range of deletion strategies suggested by a number of linguistic analyses. Although the cloze procedure has been put to extensive use, it has not been without its critics who have registered some reservations, a few of which may be worth heeding. Schlesinger (1968) advised against the continued use of the cloze technique, particularly if there was under question what he calls the "fine-grain structure of a sentence." Although his case is somewhat overdrawn, inasmuch as few researchers using the cloze technique would contend that they are concerned with the fine-grain structure of one or a set of sentences, Schlesinger cites two versions of a text and then proceeds to show in one instance what a cloze task would look like if every 2nd word were deleted:

(1) After the two friends had eaten their meal, they set out to explore the neighboring forest.
(2) The two friends, after they had eaten their meal, set out to explore the neighboring forest.
(1a) After ____ two ____ had ____ their ____, they ____ out ____ explore ____ neighboring ____.

Deleting the *same* words in (2), we obtain:

(2a) ____ two ____, after they had ____ their ____, ____ out ____ explore ____ neighboring ____. (p. 154)

Schlesinger's example admittedly is extreme—deleting every 2nd word—but his over-all point is worth noting: deletion strategies focusing on the linguistic variable of word order or sentence structure should be developed. Rather than continuing to rely almost exclusively on the grammatical elements of a sentence, conventionally defined by establishing word classes or referring to parts of speech or word categories (i.e., abstract, concrete, lexical, and structural) which are treated as isolated entities, or averaged, deletion strategies incorporating the grammatical relational nature of these word classes are worth considering as a means of refining some of the strategies currently employed. Much the same point of view has been expressed by Pike (1969) who showed proximity alone to be an inadequate index of such structural relationships as, for example, subject-verb, by citing the two following sentences which differed only in their surface structures:

The *villagers,* despite an outward acquiescence, *resisted* change.

Despite an outward acquiescence, the *villagers resisted* change. (p. 11) Whether one opts for deep or surface structure explanations, generative or interpretative semantics, or simply tries to steer clear of this complex and controversial area, it seems clear that to treat as identical even surface structure differences in language functions contributes little to an increased understanding of the phenomenon under investigation. In the present study an instance of not recognizing or acknowledging such differentiated functions can be found, for example, in the treatment of deleted nouns where no distinction is made between those nouns which serve as subject-nouns and those which serve as object nouns. Another source of speculation might be sets of lexical-deletion materials where, compared to the structural-deletion materials, a disproportionate number of main verbs has been deleted, thus perhaps posing a severe disadvantage if much weight is given to the notion that "subjects may tend to focus on the component of a sentence that is most central to its meaning or conveys the greatest amount of information, as is generally true of the predicate phrase in a statement [Snyder and Hale, 1971, p. 10]."

The difference in degree of dependence, particularly on vocabulary and associational fluency, as is indicated by the comparative high school and college data of this report, suggests that to take into account only words in isolation (vocabulary) and isolated associations, while ignoring structural, that is syntactic, relations within a sentence, can provide only partial insights into the undoubtedly complex semantic, syntactic, and associational relationships with sentences. Pike (1969) has demonstrated this point in his study of structural aspects of contextual constraint within sentences, wherein his findings reflect not only the worth of incorporating one or another scheme of structural ad semantic relationships, but also seem to demand a re-examination of the question of long-range constraint as well as the assumption that context influences predictability in only one direction—two constraint characteristics which are easily open to empirical test with a variety of cloze deletion strategies.

As for the adequacy of the criteria currently employed to develop cloze tests, some re-examination and a judicious pruning also would appear to be in order. In particular, the notions underlying the abstract/concrete distinction could bear a closer scrutiny than it generally receives. Kamman and Streeter's (1971) recent study, which compared Roger Brown's proposal that the number of known, inclusive superordinate subordinate words was the determinant of abstractness with abstractness defined in terms of empirical ratings based on accessibility to the senses, reported that "subjects appear to rate the abstractness of a generic noun in terms of the abstractness of its examplars."

> When asked whether or not ANIMAL is accessible to the senses as an object or thing, raters apparently translate ANIMAL into AN ANIMAL and accept some token image to represent the class. On the other hand, such patriarchs as CRIME, AMOUNT, and DISEASE are rated as more abstract just as their subordinates are. Possibly the words in these abstract hierarchies refer to *intrinsically relational situations* which cannot be demonstrated by finger-pointings or photographs. (pp. 303, 306. Emphasis added.)

These findings suggest some refinements in the procedures for establishing an abstract noun category, thus eliminating a few glaring inconsistencies which one occasionally finds in materials currently in use. For example, while acknowledging in specified contexts the abstract noun-ness of such items as *time, shape,* and *end* (abstract deletions in the materials used in this reported study), it is difficult to defend—even if one invokes only a rule of consistency—the labelling of *girl* as a concrete deletion and *wife* as an abstract deletion, to say nothing of allowing the words *homes* and *ornaments* as abstract deletions, and conversely, *thousands* as a concrete deletion, which, to make matters worse in this latter instance, functions as the grammatical subject of the sentence.

In sum, there seems to be a striking imbalance between the power and occasional elegance of the statistical techniques which can be and are used to assess the cloze procedure, and the hardly explointed linguistic variables which serve, at least implicitly, as one of the major underpinnings of the stimulus materials.

## REFERENCES

French, J. W., Ekstrom, R. B., & Price, L. A. *A manual for kit of reference tests for cognitive factors.* Princeotn: Educational Testing Service, 1963.

Kammann, R., & Streeter, L. Two meanings of word abstraction. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 303-306.

Miller, G. R., & Coleman, E. B. A set of 36 passages calabrated for comprehensibility. *Journal of Verbal learning and Verbal Behavior*, 1967, *6*, 851-854.

Ohnmacht, F. W., Weaver, W. W., & Kohler, E. T. Cloze and closure: a factorial study. *Journal of Psychology*, 1970, *74*, 205-217.

Pike. L. W. *Structural aspects of contextual constraint within sentences.* Research bulletin 69-16. Princeton: Educational Testing Service. 1969.

Rankin, E. F. Closure and the cloze procedure. In D. M. Wark (Ed.), *College and Adult Reading.* North Central Reading Association 3rd & 4th Yearbook. St. Paul: the Association, 1965.

Schlesinger, I. M. *Sentence structure and the reading process.* The Hague, Holland: Mouton Press, 1968

Snyder, C. W., & Hale, G. A. A new approach to the study of verbal concept learning. Educational Testing Service Research Bulletin 71-15, 1971.

Taylor, W. L. Application of cloze and entropy measures to the study of contextual constraints in samples of continuous prose. Unpublished Doctoral dissertation, University of Illinois 1957.

JOE PETERSON
University of Georgia

NAT PETERS
Oakland Schools

ED PARADIS
University of Georgia

# VALIDATION OF THE CLOZE PROCEDURE AS A MEASURE OF READABILITY WITH HIGH SCHOOL, TRADE SCHOOL, AND COLLEGE POPULATIONS

A number of studies have been designed to explore the relationship between the reader's ability to obtain closure on a mutilated passage and the more traditional determinant of readability. Rather well documented evidence (Bormuth, 1963; Rankin, 1965) has been accumulated which demonstrates that the cloze technique does measure readability and can be used for ranking a given number of passages according to relative difficulty. Recently, a few studies have yielded standards which might be used to determine whether or not a given passage is of appropriate difficulty for individual readers.

These studies, by and large, are dependent upon the standards for comprehension which Betts (1954) has determined to be appropriate for the instructional and independent reading levels, respectively 75 and 90% correct responses to questions based upon the selection.

Bormuth (1967, 1968) has reported two studies in which attempts have been made to identify a cloze percentage which corresponds to the instructional level. In his 1967 article, he reported using 73 fourth and fifth grade children to determine that multiple choice comprehension test scores of 75% and 90% corresponded with cloze scores of 38% and 50% when based on 275 word passages centered around literature, history, and science. In his 1968 study, in which he is reported to place greater confidence, he reported that 120 fourth, fifth, and sixth grade children obtained cloze scores of 44% and 57% to correspond with comprehension scores of 75% and 90% with passages taken from the Gray Oral Reading Tests. Bormuth wrote his own questions.

In an attempt to replicate Bormuth's findings, Rankin and Culhane (1969) studied the comprehension abilities of 105 fifth grade youngsters reading social studies and science articles. Their procedures differed somewhat from those used by Bormuth in that they constructed their multiple choice questions around a schema which consisted of seven types of questions, including inference questions, instead of the straightforward measurement of only those facts presented in the selection. In defense of his type of questions which are literal only, Bormuth states that his questions are lawful and therefore replicable because of their being ". . . transformations on the syntax of the sentences in the passage [1968, p. 190]." Regardless of the inconsistency in the development of questions, however, the scores obtained in the Rankin and Culhane study were within the ranges of the two Bormuth studies. Corresponding to a multiple choice comprehension score of 75% was a cloze score of 41% as compared with Bormuth's figures of 38 and 44%. The cloze score which corresponded with 90% multiple choice was 61%. This score is higher than either score (50% and 57%) reported by Bormuth. It is interesting to note that in both Bormuth's 1967 study and the Rankin and Culhane study the subjects took the cloze test over the same passage which they later read and answered questions.

These studies present data which offers a frame of reference from which one can interpret cloze results when dealing with children in upper elementary school. Although the intent of the above studies was to present a data base from which generalizations could be drawn for all levels of readers, the age range sampled seems restrictive. It is the intent of the present authors to present an exploratory study to determine the feasibility of identifying a cloze percentage which would correspond to the Betts' definition of instructional level for comprehension for high school, trade school, and college students.

## PROCEDURE

### *Subjects*

Four different groups of subjects were used. The first group was 64 high school students, three classrooms of eleventh grade pupils from a suburban area of a northern industrial city. These students were enrolled in a general curriculum and were judged to be representative average ability students of the average high school for that city. The second group was 28 high school students, eleventh graders enrolled in appliance repair in a regional vocational high school of the same northern city. The third group was 42 post-high school vocational-technical students enrolled in a medical technology training program in a multi-county southern vocational-technical school. The fourth group was 42 upper division university students who were enrolled in a reading methods course in a southern university. These students were majoring in elementary education, special education, and speech pathology.

*Data Gathering Instruments*

Two tests, one standardized and one cloze, on each subject served as the sources of data. The grade equivalent reading score of the subjects was obtained from the *Nelson-Denny Reading Test* or the *Nelson Reading Test* comprehension subtest, depending on the estimated reading ability of the individual subject.

The second sources of data were cloze tests based upon seven selections from the area of health education which were approximately 275 words in length. The decision was made to predict the grade level placement of these health passages with the Dale-Chall Readability Formula (1948) because the original validation studies had used Betts' criteria for comprehension on health education passages as the means for checking validity. That is, students reading wtih a given grade level equivalence were found able to answer correctly seventy-five percent or more of the questions asked on the health education passages which were predicted to be at their reading level.

The cloze passages were left with the first sentences intact and mutilated with an every 5th word deletion pattern. The deletions were represented with a standard 15 space line. Answers were recorded directly in the cloze passages. The passages and their Dale-Chall grade levels were: caring for teeth, below 4; the four groups of foods, 5-6; the middle ear, 7-8; composition of bone, 9-10; earache, 11-12; uses of alcohol, 13-15; psychomatic disorders, over 16.

*Data Gathering and Analysis*

The reading grade equivalent of each subject was obtained on a standardized test of reading comprehension; with most subjects, the *Nelson-Denny Reading Test* was used. Those subjects, however, whose reading grade equivalents were determined to be low enough to be in the extrapolated range (*i.e.*, below 9th grade) were retested with the *Nelson Reading Test* in order to obtain a more nearly accurate grade equivalent score. Since there were no grade equivalent scores on the Nelson-Denny given which were higher than 14+ to match the passage which was identified by formula as being over 16 in difficulty, those subjects who scored higher than the 50th per centile for grade 16 were considered to have grade 16+ proficiency in reading.

Cloze passages were assigned to subjects on the basis of a match between grade equivalent scores on the reading comprehension tests and predictions of readability on the health education passages. Subjects were informed of the purpose of the experiment and were given as long as necessary to complete the cloze passage. Exact word replacements, with allowance for decipherable misspellings, were counted as correct.

All cloze testing occurred within 10 days of the standardized reading tests. Means, standard deviations, and standard error of the mean, were computed for the groups individually and as a whole.

### RESULTS AND DISCUSSION

The results for the four groups of subjects do not appear to be in agreement with those found by Bormuth or Rankin and Culhane. Further examination of mean cloze scores reveals an inconsistency among the groups of subjects measured. The mean scores for the general high school and vocational high school subjects were considerably greater than mean scores for the adult vocational-technical students and the university students.

Table 1
Results of Cloze and Reading Tests

| Subjects | Mean Cloze Percentages (SD) | Mean Reading Grade Levels (SD) |
|---|---|---|
| Total Group (N=176) | 50 (14) | |
| General H.S. (N=46) | 57 (12) | 7.3 (1.3) |
| Vocational H.S. (N=28) | 57 (16) | 4.9 (1.7) |
| Adult-Vocational-Technical (N=42) | 43 (12) | 11.8 (1.7) |
| University (N=42) | 44 (10) | 14.3 (1.7) |

The mean scores of the two high school groups greatly exceeded the score for the instructional level found by Bormuth or Rankin and Culhane and in fact approached the independent level reported by these investigators. This similarity to the independent level combined with the vast differences between the high school subjects and the adult subjects suggested a search for supporting data. The reading comprehension subtest scores indicate that the adult vocational-technical subjects and the university subjects have mean reading scores at a level one would expect. The adult vocational technical subjects were in a medical technology program and judged by their teachers to be average high school graduates. The university subjects were beginning their third year in an elementary curriculum and while the mean score is not at grade level, it is not unusually low.

The mean reading scores for the two high school groups are lower than one would expect. The general high school group was drawn from a school judged as average for a midwestern urban area. A typical group of eleventh grade students from an average school would be expected by definition to have a mean grade level score of 11.0. A mean score of 7.3, however, does not appear to be an accurate measure of a typical group. The mean score for the vocational high school subjects was also low. While the vocational high school students may not typically be expected to read at grade level, a mean score of 4.9 did not seem right.

The inconsistency of reading achievement scores among the four groups of subjects seemed to warrant a further search for information. No additional information was available for the general high school subjects, the adult vocational-technical subjects, or the university subjects. However, the cumulative folders for a selected group of vocational high school subjects were available.

Accordingly, the cumulative folders for the eight vocational high school subjects receiving the lowest reading achievement scores were examined. *Gates-MacGinitie Reading Tests* scores from the fall of the school year indicated higher reading abilities than measured on the *Nelson Reading Test*. The mean grade level difference on the comprehension subtests was 2.9 years. While this information is from a small number of selected subjects and may involve regression to the mean, it does seem to question the amount of confidence that can be placed in these *Nelson Reading Test* scores.

Additional investigation revealed that many of the high school subjects had received several standardized tests just prior to receiving the *Nelson Reading Test*. It is possible these subjects were tested out and put forth only a partial effort. If this was the case, one can only hypothesize that the cloze task intrigued the subjects and they put forth a maximum effort with a passage well below the instructional level. Thus, the result would be an unusually high schore on the cloze task, approaching the independent levels as defined by Bormuth or Rankin and Culhane.

A second possible cause for the high school subjects' success could be their ability with the cloze procedure. If the reading grade levels are correct, then these subjects have likely been functioning in reading material written beyond their instructional level. Students continually faced with material in which they can identify relatively few words may develop advanced techniques in guessing words from the context of the passage. This guessing skill could transfer to a cloze task. Thus, these subjects might be paired with passages written at their instructional level but advanced skill in use of context clues enabled them to score unusually high on the cloze task. The data from the high school subjects seems to raise more questions than it answers. Additional information seems necessary before any conclusions can be made.

The results from the adult vocational-technical subjects and the university subjects differ from those of the high school subjects, while they are similar to those found by Bormuth and Rankin and Culhane. The criteria determined by these researchers with elementary school children seem appropriate for use with adults.

## SUMMARY AND CONCLUSIONS

The purpose of the present study was to determine whether or not the cloze percentage score which Bormuth has identified as corresponding to the instructional level for intermediate grade pupils is applicable to more sophisticated students. Accordingly, passages on health education, the readability of which had been determined with the Dale-Chall readability formula, were mutilated with an every 5th word deletion pattern and given to 176 high school and post-high school students on the basis of their measured reading abilities.

The results were not directly comparable with those obtained by Bormuth. The mean scores for the group as a whole were 50% as compared with Bormuth's finding of 44%.

The data were examined further because of this discrepancy. Mean scores for the high school and vocational high school were 57% while those for the vocational-technical and university were 43% and 44%. This unexpected range of scores and their subsequent grouping according to the age of the pupils remained unexplained. A sampling of the cumulative records, available only on the high school age students, for previous reading scores on other standardied tests yielded scores which were higher than those obtained in the study. It was hypothesized that the cloze material given these students may have been at a difficulty level below their actual reading level. Directly related to this hypothesis is another, that the cloze scores of these groups would have approximated those of the post-high school groups had the reading levels been more nearly accurately obtained. These hypotheses are, however, yet unproved, and the possibility exists that the scores were accurate and the data reflect a superior cloze ability on the part of the general and vocational high school students. Additional data is warranted before conclusion can be drawn.

No major discrepancies were observed or inferred with the adult groups. It does seem appropriate to conclude that it is possible to use Bormuth's criteria for determining the instructional level with adults.

## REFERENCES

Betts, E. A. *Foundations of Reading Instruction*. New York. American Book, 1954.

Bormuth, J. R. Cloze as a measure of readability. *Yearbook of the International Reading Association*. Newark, Delaware: International Reading Association, 1963, *8*, 131-134.

Bormuth, J. R. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 1967, *10*, 291-299.

Bormuth, J. R. Cloze test readability: criterion reference scores. Journal of Educational Measurement, 1968, *5*, 189-196.

Dale, E., & Chall, J. S. A formula for predicting readability. *Educational Research Bulletin*, 1948, *27*, 11-20, 28.

Rankin, E. F., Jr. The cloze procedure—a survey of research. In E. L. Thurstone & L. E. Hafner (Eds), *The philosophical and sociological basis of reading*. National Reading Conference 14th Yearbook. Milwaukee: The Conference, 1965. Pp. 133-148.

Rankin, E. F., & Culhane, J. W. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 1969, *13*, 193-198.

EARL F. RANKIN
TOM J. KEHLE
University of Kentucky

# A COMPARISON OF THE READING PERFORMANCE OF COLLEGE STUDENTS WITH CONVENTIONAL VERSUS NEGATIVE INTERNAL (INTRA-ARTICLE) READING FLEXIBILITY

There are many concepts of "reading flexibility" and numerous ways of measuring various types of reading flexibility. A recent survey of the literature by Rankin (1972b) provides abundant evidence for this obvious assertion. However, most conceptualizations of reading flexibility include the assumption that the flexible reader will adjust his reading approach and rate by slowing down for more difficult purposes and/or materials and speeding up for easier reading purposes and/or materials. In an important study of the relative influence of material and purpose on the reading rate of 601 college freshmen, Letson (1959) found a group of 96 students who reversed their rates by reading the more difficult of two passages at a faster rate than the easier passage. Similarly, he found a group of 203 students out of the total group who slowed down when instructed to read one passage as rapidly as possible and speeded up when instructed to read another passage of equivalent difficulty for complete mastery of ideas and details. These students were termed negative groups as contrasted with the positive groups whose reading flexibility was in accordance with conventional expectations. A comparison of the comprehension scores of the negative versus positive groups showed that there was little difference in the mean comprehension scores for easy materials. However, the positive group attained a higher comprehension percentage score of 62% on the difficult passage as contrasted to a comparable score of 54% for the negative group. Although no tests of significance were given, Letson concluded that flexible readers (in the conventional sense) read difficult materials with better understanding than readers whose flexibility was characterized by changes in rate which are opposite to most conceptions of reading flexibility. Interestingly, he found no difference between comprehension scores for these two flexibility groups as a function of different assigned purposes while reading materials of equivalent difficulty. In fact, he observed that directions to read for complete mastery tend to make some students "sit up straight and read with more attention—and often with greater speed [p. 240]."

These findings suggest a new dimension of reading flexibility which might be termed negative flexibility in contrast to conventional flexibility. Although teachers who use standardized tests of reading flexibility frequently have found individual students who display negative flexibility," we have not located other investigations of this phenomenon in the literature.

Rankin and Hess (1970) and Rankin (1972a) have devised a technique for measuring what McCracken (1965) has termed internal flexibility. This term refers to the adjustment of rate and approach within the sentences and paragraphs which make up a total passage. The previously mentioned measurement technique results in an individual flexibility coefficient for each reader designating the strength and direction of the relationship between the reader's rate of reading for successive word segments which make up a total passage and the readability of these segments. The technique is ideally suited for the investigation of negative flexibility.

The purposes of the present study were to compare:

A. Reading comprehension scores of 'poor' college students who may be characterized as having conventional versus negative internal reading flexibility, both before and after training.

B. Reading rate scores of 'poor' college students who may be characterized as having conventional versus negative internal reading flexibility, both before and after training.

## PROCEDURE

This investigation consisted of a re-analysis of the data from a previous study by Rankin (1972). Subjects were 255 college students enrolled in a reading improvement program who were selected on the basis of having attained scores below the 33rd percentile on the *Cooperative English Tests: Reading Comprehension*. Most of these students were freshmen. Both before and after a one-semester course in reading improvement, during which a great deal of training was devoted to the enhancement of conventional reading flexibility, their internal reading flexibility was based on alternate forms of the Story Comprehension Sub-test of the *Diagnostic Reading Tests: Survey Section (DRT)*. These passages were written at a rather easy level, because the test was designed to be taken by students from the seventh grade to the college level. Form A was administered before training and Form C after training. After receiving a practice exercise to adapt them to the testing procedure, the subjects read the approximately 2,000 word passage with instructions to underline the word they were reading each time the experimenter said "mark." This signal was given every 15 seconds. The assigned purpose for reading was to "Read as rapidly as you can and still understand what you read." After reading the article, each subject attempted to answer a 20-item test based upon materials in the passage without looking back at the passage. Intra-article reading rates for each subject were computed by determining the number of 15 second intervals it took him to read each successive 100-word segment of the total passage. The readability of each 100-word segment was determined by using the cloze procedure on these passages with a comparable group of subjects. An individual flexibility coefficient was computed for each reader by computing a Pearson correlation between the distribution of his reading rates for each successive 100-word segment and the distribution of cloze readability scores for these segments. It should be noted that a negative correlation indicates conventional flexibility because fast reading is indicated by a small number of 15 second intervals for a 100-word segment, and high readability (i.e., easy reading) is reflected in a high cloze score for a given passage. Conversely, negative flexibility is denoted by positive flexibility coefficients.

## RESULTS

The distribution of flexibility coefficients based upon previous findings by Rankin (1971) is given in Figure I. Figure I dramatically illustrates the extent of both conventional and negative flexibility coefficients in this student group. The results of this investigation relative to the previously stated purposes are given in Table 1.

Before training, 193 readers displayed conventional flexibility (i.e., negative correlations), and 62 readers showed negative flexibility (i.e., positive correlations). After training, comparable figures were 228 and 27, respectively. It would appear that training may have created more conventional reading flexibility in this group. In the absence of a control group, this latter statement is, of course, only a supposition.

An analysis of pre-training results (based on Form A, *DRT*) reveals that the mean flexibility coefficient for the conventional group was −.26 as contrasted to a mean flexibility coefficient of .16 for the negative group. Although these mean flexibility coefficients were low, it should be remembered that some previous research has indicated that poor readers display less reading flexibility than good readers. Even so, the mean coefficient of −.26 is statistically significant for this large group. A zero correlation would indicate no flexibility. The t-test

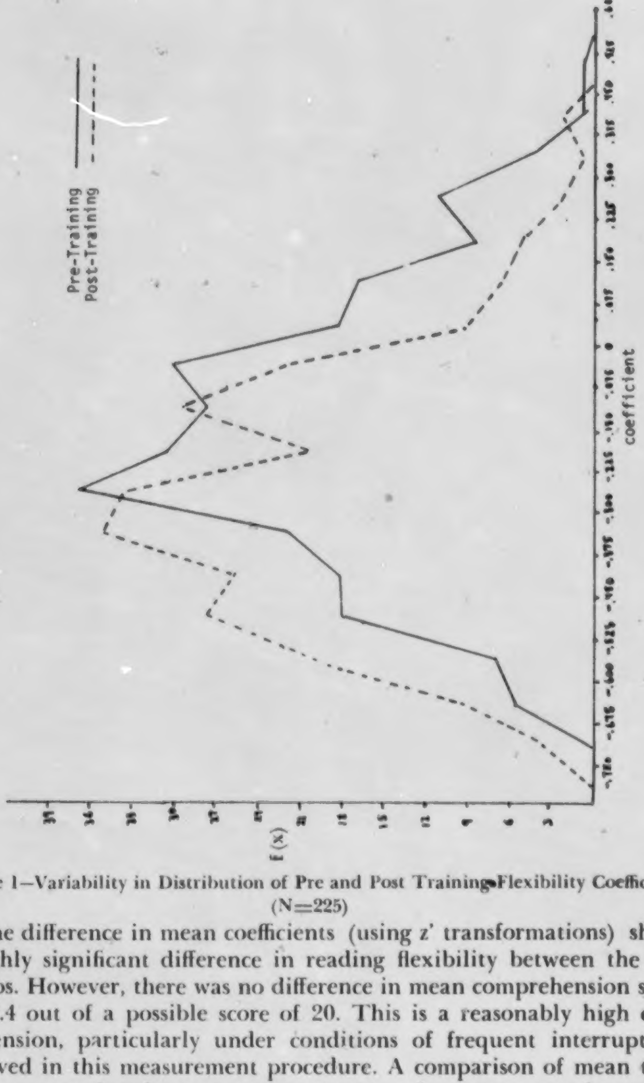


Figure 1—Variability in Distribution of Pre and Post Training Flexibility Coefficients (N=225)

for the difference in mean coefficients (using z' transformations) shows a highly significant difference in reading flexibility between the two groups. However, there was no difference in mean comprehension score of 15.4 out of a possible score of 20. This is a reasonably high comprehension, particularly under conditions of frequent interruptions involved in this measurement procedure. A comparison of mean total reading rates for the two groups (expressed as mean number of 15 second intervals for the total passage) also showed no significant difference. Expressed as words per minute (WPM), the conventionally flexible group read at 258 WPM and the negatively flexible group read at 251 WPM. Therefore, it can be concluded that although there was a highly significant difference in reading flexibility between the two

Table 1

Mean Flexibility Coefficients, Comprehension Scores and Reading Rate Scores Before and After Training for Subjects Wtih Convention versus Negative Reading Flexibility

| Type of Flexibility | | Pre-Training | Post-Training | |
|---|---|---|---|---|
| | | Mean Flexibility Coefficients | | |
| Negative | | (n=62)<br>X = .16<br>s = .11 | (n=27)<br>X = .14<br>s = .11 | |
| Conventional | | (n=193)<br>X = −.26[1]<br>s = .16 | (n=228)<br>X = −.32[1]<br>s = .18 | |
| | | Mean z' Transformations of Flexibility r's | | |
| Negative | t = 17.99<br>p<.001 | (n=62)<br>X = .16<br>s = .12 | (n=27)<br>X = .14<br>s = .12 | t = 11.54<br>p<.001 |
| Conventional | | (n=193)<br>X = −.28<br>s = .19 | (n=228)<br>X = −.35<br>s = .22 | |
| | | Mean Comprehension | | |
| Negative | t = −.13<br>p>.05 | (n=62)<br>X = 15.4<br>s = 2.61 | (n=27)<br>X = 13.7<br>s = 2.67 | t = −1.86<br>p>.05 |
| Conventional | | (n=193)<br>X = 15.4<br>s = 2.66 | (n=228)<br>X = 14.7<br>s = 2.53 | |
| | | Mean Rate[2] | | |
| Negative | t = 1.13<br>p>.05 | (n=62)<br>X = 31.89<br>s = 5.39 | (n=27)<br>X = 24.21<br>s = 5.70 | t = −3.00<br>p<.001 |
| Conventional | | (n=193)<br>X = 30.94<br>s = 5.82 | (n=228)<br>X = 27.64<br>s = 5.60 | |

[1] p<.01

[2] Rate—Mean number of 15 second intervals taken to read passage

groups of poor college readers, this variable bore no relationship to either reading comprehension or rate before training.

Post-training results were somewhat different. Here again, the difference in mean flexibility coefficients for the two groups was highly significant. Changes in the number of subjects before and after training displaying conventional versus negative reading flexibility were in the expected direction, as were the mean flexibility coefficients of both groups. It is important to note that the notoriously non-equivalent forms of the *DRT* resulted in the post-test (i.e., Form C) being more difficult than the pre-test (i.e., Form A). This statement is based not only on the finding that both groups made lower comprehension scores on the post-test, but also upon previous research by Coates (1968) which revealed that Form C is more difficult than Form A on both comprehension and vocabulary sub-tests, as well as the total test score. In any case, although the difference in mean comprehension scores after training favored the conventionally flexible group, this difference was not quite statistically significant. However, a difference of only one point would be of little educational significance. Letson (1959) found no difference in comprehension between his positive and negative groups on easy materials, but found significant differences on difficult materials.

Perhaps the most interesting finding in this study was the comparison of the total passage mean reading rates for these two groups before and after training. As previously described, there was no significant difference before training between the mean overall reading rates for the two groups. However, after training, the difference in mean rates was significant. Although both groups improved in rate when comparing pre- and post-training rates, the important finding here is that the group reading with conventional flexibility read with a significantly slower reading rate after training than the negative group. Expressed in words per minute, the conventionally flexible group read the post-training passage at 289 WPM as compared to a 330 WPM rate for the negative flexibility group.

If it may be assumed that the reading passage on the post-test was more difficult than the passage on the pre-test, then the slower rate of the conventionally flexible group after training (in contrast to the negatively flexible group) might reflect a desirable outcome of the reading improvement course. Both groups, of course, improved in rate during the semester. On the other hand, if the passages were written at an equivalent readability level, but the test questions on the post-test were more difficult than the items on the pre-test, then the interpretation of these findings would be somewhat different. Under this assumption, it may be that these poor readers who had been trained to read with conventional flexibility became more sensitive to their ability to process information at various speeds and, hence, read at a slower overall speed commensurate with their generally limited reading ability. Also, they probably paid more attention to the more difficult test questions on the post-test than the other group. Their comprehension loss from pre-to post-test comprehension was only .7 of a point as contrasted to a loss of 1.7 points for the negatively flexible group.

## SUMMARY

Several limitations of this investigation should be mentioned. First, the subjects were selected on the basis of low reading achievement. Second, the materials were not only rather easy, but also, the distribution of cloze scores for successive word segments in the reading passages displayed a limited range. Both of these factors would probably tend to reduce the degree of intra-article reading flexibility. Within these limitations, two general interpretations might follow from the results. One interpretation might be that training for improvement in conventional flexibility results in a desirable outcome for poor readers to make realistic adjustments in reading rates as a function of changes in difficulty of reading passages and/or test items. Another possible interpretation is that: The game is not worth the candle. That is, these students who were reading with varying degrees of conventional flexibility were reading, after training, at a rate which was 41 WPM slower than the negatively flexible group and achieved only a small difference of one point in the comprehension score. In terms of efficiency, one might question if the outcome was worth the price.

The latter interpretation raises a fundamental question about the importance of various conventional concepts of reading flexibility. Levin (1968) and others have found only very small correlations between measurements of different types of reading flexibility and reading achievement as measured by standardized tests, intelligence, and school performance. One might question the significance of the whole concept of reading flexibility. On the other hand, as Rankin (1972b) has pointed out, there are many deficiencies in measurement procedures for reading flexibilities. Such deficiencies may account for many negative and conflicting findings regarding correlates of reading flexibility.

It may be that research utilizing the measurement of internal flexibility may shed some light on the nature and correlates of reading flexibility—both conventional and negative. Future studies of internal flexibility should utilize normal groups of readers at various age levels, materials written at different degrees of difficulty and containing various amounts of heterogeneity with respect to fluctuations in intra-article readabilities of word segments. In the meantime, a good deal of speculation and advocacy of the value of reading flexibility which has become a part of our conventional wisdom, should be curtailed until a substantial body of solid evidence is gathered to substantiate our claims. At present, we cannot really be sure that what has been called negative flexibility in this study is after all an undesirable reading technique, at least for all readers, materials, and reading purposes.

## REFERENCES

Coates, L. F. The enigma of the survey section of the Diagnostic Reading Tests. In G. B. Schick & M. M. May (Eds.), *Multidisciplinary aspects of college-adult reading*. National Reading Conference 17th Yearbook. Milwaukee: the Conference, 1968. Pp. 70-78.

Letson, C. T. The relative influence of material and purpose on reading rates. *Journal of Educational Research*, 1969, *52*, 238-240.

Levin, D. J. The flexibility of reading rate. *International Reading Association Proceedings*, 1968, *12* (1), 596-603.

McCracken, R. A. Internal versus external flexibility of reading rate. *Journal of Reading*, 1965, *8*, 208-209.

Rankin, E. F., & Hess, A. K. The measurement of internal (intra-article) reading flexibility. In G. B. Schnick & M. M. May (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook, Vol. 1. Milwaukee: the Conference, 1970. Pp. 254-262.

Rankin, E. F. How flexibly do we read? *Journal of Reading Behavior*, 1971, *3*, 34-38.

Rankin, E. F. The measurement of reading flexibility. ERIC/CRIER Reading Information Series: Where Do We Go?, In press.

NED H. RATEKIN
University of Northern Iowa

# THE ADEQUACY OF THE CLOZE IN MEASURING COMPREHENSION OF DIFFERENT LOGICAL PATTERNS

This study explores the extent to which multiple-choice and cloze tests agree in their measurement of comprehension of material presented in two logical patterns, induction and deduction.

Considerable effort has been expended over the last few years in validating the cloze technique of testing reading comprehension. This validation for the most part has compared students' performance on cloze tests with their performance on standardized reading tests, or in some instances, on prepared multiple-choice tests. The high relationships found in such studies as Culhane (1970), Rankin (1970) and Bormuth (1967) have generally been accepted to indicate that cloze tests approximate the general reading comprehension skill as measured in standardized tests.

Weaver and Kingston (1963), among others, have claimed that cloze testing actually samples a different type of reading behavior than does multiple-choice testing. Bormuth's response (1969) was to present evidence of a general reading ability that is represented equally by constructed multiple-choice tests and cloze tests over five sub-skills. He was unable to identify on the basis of student comprehension scores any factors or skills that were separate from this general factor on either the cloze or multiple choice tests.

This study is an attempt to look again at the assumption that cloze and multiple-choice tests equally reflect a general reading performance. Although predictions are high, cloze tests do not, of course, predict exact performance on other reading tests. This difference in variance may well represent a strength or advantage to cloze testing rather than indication of lack of validity. It is on an intuitive level very tempting to view the cloze test, which requires close scrutiny of passages, as measuring some different processes, or skills than those measured by standardized multiple-choice tests.

There is a long history of frustration and difficulty in isolating specific reading skills for analysis or for instruction in the skill. Mosberg's work (1969) is an interesting example of attempts to categorize knowledge in this area. A variety of categories of factors have been suggested, such as: those related to organizational pattern, linguistic units related to comprehension, following patterns of thinking which result in adequate analysis of material, and the relationship of cognitive development to specific aspects of reading performance. These attempts to characterize skills in reading have identified factors in the structure and organization of the materials read, or in the structure and organization of the menatl processes of the reader as factors in the comprehension process.

There is a paritcular subset of these factors which refer simultaneously to the structure of ideas in print and to the structure of particular cognitive skills. For example, reading for chronological or sequential order is an application of seriation skills; making a generalization from stated facts is a specific type of inductive reasoning; and following a cause and effect sequence is a type of conditional reasoning. The ability by the reader to identify and follow these reasoning patterns may underlie other skills of comprehension such as those labeled interpretation, evaluation, and analysis.

A pair of general skills which fall in this category are the reasoning skills of induction and deduction related to patterns of idea development in written materials. The primary questions asked int his study center on the relative ability of cloze and multiple-choice tests to measure a reader's ability to comprehend materials in these two logical patterns.

For this study, inductive reasoning is identified as the process of forming generalizations or drawing conclusions on the basis of facts, reasons or illustrations, by analogy, or through the use of comparison and contrast. An idea in print which presents a generalization by such evidence is identified as an inductive logical pattern. When this pattern appears in print, the generalization may either precede or follow the evidence. The logic in this type of reasoning is based on probability.

Deductive reasoning in this study is defined as a process of drawing conclusions on the basis of logical necessity, or in following a form that implies such a conclusion. The forms of concern are the traditional syllogistic forms, mathematical reasoning, transpositional reasoning and conditional reasoning. As in the inductive pattern, natural language passages identified as deductive do not necessarily follow the formal order of discovery. In a paragraph or longer passage, the conclusion may well precede the logical evidence. The premises on which the conclusion is based then appear as details of the passage.

There is adequate evidence that these two patterns of reasoning exist and develop as unique abilities. Wetherick (1969) and Wetherick and Freeland (1969) found inductive inference to be a general capacity separate from any specific content, and that this capacity developes in two stages. In the first stage the person learns to identify the subset of cues in common instances. In the second stage, the person learns to identify the subset of cues not present in other instances. He found this last skill to be acquired last in childhood and lost first in old age, implying to him that a short-term memory factor affects the second stage.

According to Kofsky (1966) learning the induction process, in Piagetian terms, is learning resemblance sorting, consistency in sorting, and exhaustive classification, while the deductive process develops through learning multiple class memberships, horizontal classification and hierarchical reclassification. Roberge (1970) also identifies deductive reasoning with the mental operations of class reasoning and conditional reasoning. He has illustrated developmental improvement in deductive reasoning over grades 4-10.

Although these patterns of reasoning are apparently different in form and development, there is little evidence to compare them in difficulty. Since class reasoning and conditional reasoning grow out of inductive experience, it may be assumed that deduction begins to develop after some inductive skills have been established.

It is assumed in this study that ability to follow these two logical patterns in written materials would be a skill related to successful reading comprehension. In fact, reading experts in their texts and other writings frequently refer to the need to develop skill in discriminating logical patterns. (See for evample, Jordan, 1967; Karlin, 1971; Herber, 1971.)

Despite the importance of this skill, standardized tests provide minimum opportunity to evaluate student abilities in a range of these patterns. Table 1 presents an analysis of the provision for inductive and deductive reasoning in six well-known standardized reading tests. The reading passages in the comprehension portions of these tests have been designated as inductive or deductive. Narrative and descriptive material is included in the inductive category.

An extremely high incidence of inductive items was found in these tests. Part of the difference found between results of cloze and multiple-choice testing may be due to skills, such as deductive reasoning, which are not sampled on standardized tests. In such a case, the cloze test with the possibility for selecting a wider range of patterns, may give a

Table 1

Per Cent of Inductive and Deductive Passages and Related Questions in Six Standardized Reading Tests

| Reading Test | Inductive | | Deductive | |
|---|---|---|---|---|
| | Passages | Questions | Passages | Questions |
| **Coop English, Form 1A** | 100 | 100 | 0 | 0 |
| **Diagnostic Reading Test: Survey Form B** | 83 | 94 | 17 | 6 |
| **Gates-McGinitie, Survey E. Form 1M** | 76 | 75 | 24 | 25 |
| **Iowa Silent Reading Test, Form Am** | 100 | 100 | 0 | 0 |
| **Nelson Denny Reading Test, Form A** | 100 | 100 | 0 | 0 |
| **Stanford Diagnostic Reading Test, Level 1** | 68 | 78 | 32 | 22 |

more adequate measure of this ability for an individual reader.

The purpose of this study was to discover if any difference between comprehension of material presented inductively and comprehension of material presented deductively is equally reflected by scores on multiple-choice and on cloze tests.

Specifically, the following two null hypotheses are tested in the study.

1. There is no difference between comprehension scores of students reading materials identified as inductive and materials identified as deductive
   a. As measured by a multiple-choice test.
   b. As measured by a cloze test.
2. There is no difference between correlations found when comparing inductive and deductive scores from multiple-choice tests and those scores from cloze tests.

To test these hypotheses the test Comprehension of Logical Patterns, was given to 33 college freshmen and sophomores who were among those enrolled in a community college reading improvement program. Students in the program were advised to enroll on the basis of their scoring below the 40th percentile on the Diagnostic Reading Test.

The Logical Patterns Test, prepared especially for this study, is composed of two sections, one a series of multiple-choice tests and the other a series of cloze tests. Each section of the test is divided into two parts. In one part the passages are presented inductively, and in the second part the passages follow a deductive pattern. These passages were selected from three different sources. Selections of approximately the same length were included in each part of the test from each of the three sources. The difficulty level of the passages was identified by the Fry scale as grade 10, grade 12 and college (difficult, near the end of the chart). Multiple-choice questions were constructed for the first two parts of the test. Each of the 40 items offered four foils, and required relating ideas in the passage.

The cloze tests appeared with every 5th word deleted and required exact replacement for correct responses. Special answer sheets were prepared for both the multiple choice and cloze items. The test was administered in two sessions.

Table 2 presents a design of the test which yields four subtest scores: MC-I, MC-D, Cloze-I, Cloze-D.

Table 2

Test I—Comprehension of Logical Patterns

| | Cloze | | Multiple-Choice | |
|---|---|---|---|---|
| | Inductive | Deductive | Inductive | Deductive |
| Sources | 1 2 3 | 1 2 3 | 1 2 3 | 1 2 3 |
| Tests<br>(Scores) | MC-I<br>(20) | MC-D<br>(20) | Cloze-I<br>(139) | Cloze-D<br>(145) |

The data was analyzed by t-tests of the significance of difference between correlated means, and by comparing correlations in the matrix formed from the four test scores.

## RESULTS AND DISCUSSION

The first hypothesis requires a comparison of the difficulty students found in reading the two logical patterns. Table 3 presents the results of t-tests for the significance of difference between correlated means. These scores are expressed in percent of correct responses.

Table 3

t-Tests of Difference Between Correlated Means on Multiple-Choice and Cloze Tests of Inductive and Deductive Passages

| Test form | Comprehension Means (S.D.) | | t |
|---|---|---|---|
| | Inductive | Deductive | |
| Multiple-Choice | 51(13) | 45(12) | 2.218* |
| Cloze | 30(10) | 23(9) | 5.742** |

*p<.025
**p<.001

In relation to the accepted values of 75 per cent comprehension on multiple-choice and 40 per cent on cloze as levels of frustration, the passages on this test must be viewed as quite difficult for the students. The Inductive-deductive difference on both the multiple-choice and the cloze tests are large enough to be significant. The deductive passages are consistently more difficult.

Therefore, the hypothesis that there is no difference between comprehension of inductive and deductive passages must be rejected on the basis of scores from both tests. The somewhat larger difference found on the cloze test implies that it may be more sensitive to differences in the difficulty students found in understanding indutcive and deductive passages.

The second hypothesis requires a comparison of correlations among the subtest scores to determine if either test is more sensitive to individual differences among students in their ability to comprehend the two types of material. Table 4 presents the matrix of correlations found among the four test scores.

Table 4
Correlations Among Multiple-choice and Cloze Subtests of Inductive and Deductive Comprehension

| Test | MC-1 | MC-2 | Cloze-1 | Cloze-2 |
|---|---|---|---|---|
| MC-1 | — | .24 | .40 | .20 |
| MC-2 |  | — | .55 | .38 |
| Cloze-1 |  |  | — | .80 |
| Cloze-2 |  |  |  | — |

In general these correlations fail to reflect the degree of relationship expected between tests of the same form or pattern. Beyond the possibility of low test reliability, these lower correlations may be due to the fact that each of the subtests differ from each other in at least one dimension, pattern or test form, and two of them differ in both dimensions.

When test form is held constant, the cloze test correlates highest across the two patterns, with a correlation of .80. On the multiple-choice tests the correlation is .24. The two tests did not produce similar correlations between scores on inductive and scores on deductive materials. The relatively high correlation of .80 shows a strong relationship between comprehension of the two logical patterns when measured on the cloze test. But when measured by multiple choice items, the correlation .24, does not reach a level beyond chance. In spite of the necessary reservations due to test reliability and item validity, it seems that multiple-choice tests, prepared for that purpose do discriminate between student's comprehension of inductive and deductive patterns to a greater extent than do cloze tests.

Therefore, the hypothesis that there is no difference between the correlations found when comparing inductive and deductive scores on the two tests must be rejected.

## CONCLUSIONS

In this study the scores from a multiple-choice test prepared to measure comprehension of inductive and deductive logical patterns were compared to scores from cloze tests of the same patterns. Results on both tests indicated that the deductive patterns were significantly more difficult to understand than were the inductive passages. Also, the low correlation found on the multiple-choice test between comprehension of deductive and inductive patterns, compared to the high correlation of the two skills on the cloze tests, suggests that the multiple-choice tests prepared to measure comprehension of these patterns does a better job of discriminating differences in an individual reader's ability in the two areas.

## REFERENCES

Bormuth, J. R. Comparable cloze and multiple choice test scores. *Journal of Reading*, 1967, *10*, 291-299.

Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension ability. *Reading Research Quarterly*, 1969, *4*, 358-365.

Culhane, J. W. Cloze procedure and comprehension. *Reading Teacher*, 1970, *23*, 410-413.

Darnell, D. K. The relationship between sentence order and comprehension. *Speech Monograph*, 1963, *30*, 97-100.

Herber, H. L. *Teaching reading in the content areas.* Englewood Cliffs: Prentice-Hall, 1971.

Jordan, L. V. Building the readiness program. *Mental Retardation*, 1967, *5*, 28-31.

Karlin, R. *Teaching elementary reading*. New York: Harcourt, Brace, Janovich, 1971.

Kofsky, E. A scalogram study of classification of development. *Child Development*, 1966, *37*, 191-204.

Mosberg, L., & Shima, F. *Comprehension of connected discourse*. Inglewood: Southwest Regional Educational Laboratory, 1969.

Rankin, E. F., & Culhane, J. W. Comparable cloze and multiple choice comprehension test scores. *Journal of Reading*, 1970, *13*, 410-413.

Roberge, J. J. A study of children's ability to reason with basic principals of deductive reasoning. *American Educational Research Journal*, 1970, *7*, 583-596.

Wetherick, N. E. Structure and content in concept attainment: a full scale study. *British Journal of Psychology*, 1969, *60*, 347-350.

Wetherick, N. E., & Freeland, C. Inductive thinking in children of primary school age. *British Journal of Psychology*, 1969, *60*, 11-16.

Weaver, W. W., & Kingston, A. J. A factor analysis of cloze procedure and other measures of readability. *Journal of Communication*, 1963, *13*, 252-261.

WILSON L. TAYLOR
Special Mental Health Research, NIMH
Saint Elizabeths Hospital

# RELATIVE INFLUENCES OF PRECEDING, FOLLOWING, AND SURROUNDING CONTEXTS ON CLOZE PERFORMANCE

This study explored the relative influences of "preceding," "following," and "surrounding" contexts on cloze responses to blanks in samples of prose. It asked if subjects would perform differently—i.e., guess more missing words correctly *or* interagree more consistently among themselves—if they saw the five words which preceded a blank instead of the five which followed it. And how would those two sets of performances compare with responses to both contexts presented together, with each blank "surrounded" by 10 words, five before and five after?

## PROCEDURE

### *Stimulus Materials*

Forty-five prose fragments, each 11 words long—with the middle (sixth) word to be replaced by a blank and guessed from context—were systematically drawn from a short story.

Every ninth line in the printed text was marked. The first full word in each marked line became the beginning of a series of 11 consecutive words unless the middle word was judged to be a wholly unpredictable proper noun, in which case the 11-word series began with the second full word, or the third if necessary. This process yielded many fragments made up of different but neighboring sentences. Punctuation marks were preserved. All blanks were of a standard length.

Three cloze versions of each fragment were prepared. The surrounding context version (Sur) exhibited five words on each side of the blank. The preceding version (Pre) had five words before the blank, and the following version (Fol) showed the blank and the five words after it.

The 135 stimulus items (three versions each of 45 fragments) were systematically sorted into three test assemblies, i.e.:

| Fragments Assigned: Sets of 15 Out of 45 | Assemblies | | | |
|---|---|---|---|---|
| | A | B | C | |
| #1, #4, #7, . . ., #43 | Pre | Sur | Fol | Context |
| #2, #5, #8, . . ., #44 | Fol | Pre | Sur | Versions |
| #3, #6, #9, . . ., #45 | Sur | Fol | Pre | Assigned |

All fragments, 15 in each context version, were represented in each assembly. In aggregate, one set of three different assemblies presented every version of every fragment exactly once.

### ADMINISTRATION

Each subject was asked to make one guess about the identity of each missing word. Cloze responses were elicited from 18 journalism students, six to each assembly, at the University of Illinois, Urbana. (That was in 1956, but the analysis presented here was not completed until 1971, after development of rational methods for quantifying degrees of response-interagreement.)

#### *Scoring Methods*

Every one of the 810 cloze responses (45 fragments times 18 subjects) was scored as like or unlike (1) the original word deleted and (2) the responses of other subjects. Any two words were considered alike if they were obviously the same—even if they differed in spelling or capitalization. Not counted as alike, however, were synonyms or word forms which differed in number, tense, case, or gender.

Those guesses which matched the original deleted words were summed to produce match-original (MO) scores (in the past, frequently referred to simply as "cloze scores") for each subject, group, context version, and fragment.

Degrees of interagreement among responses (without regard for the identities of the original words) were scored in three ways. Within every group of six subjects and for every fragment, each individual received a frequency of interagreement (fIa) score, i.e., the number of other subjects with whom he agreed. Then the individual fIa scores were summed to obtain an overall subject-interagreement (SI) score for the group. Also, for each fragment, the responses to Pre and Fol contexts were compared to Sur responses to obtain Like Sur measures for Pre and Fol. This measure was the number of Pre (or Fol) responses which could be paired with Sur responses.

All these methods of quantification are illustrated in Table 1 along with all of the individual responses to Fragment #3. Similar full information on all 45 fragments was included in the convention paper, Taylor (1971), from which this report is derived.

Table 1

Responses to Fragment #3; Methods of Scoring Them

| Surrounding Context | | | Preceding Context | | | Following Context | | |
|---|---|---|---|---|---|---|---|---|
| . . . as a farm hand, sell _____ horse and embark on an . . . | | | . . . as a farm hand, sell _____ . . . | | | . . . _____ horse and embark on an . . . | | |
| Ss | Responses | fIa | Ss | Responses | fIa | Ss | Responses | fIa |
| A1 | HIS | 2 | C1 | apples | 0 | B1 | the | 2 |
| A2 | HIS | 2 | C2 | eggs | 0 | B2 | a | 1 |
| A3 | HIS | 2 | C3 | chickens | 0 | B3 | the | 2 |
| A4 | a | 0 | C4 | milk | 0 | B4 | the | 2 |
| A5 | the | 1 | C5 | oats | 0 | B5 | a | 1 |
| A6 | the | 1 | C6 | labor | 0 | B6 | ride | 0 |
| MO: 3 SI: 8 | | | MO: 0 SI: 0 | | | MO: 0 SI: 8 | | |
| | | | Like Sur: 0 | | | Like Sur: 3 | | |
| | | | Like Sur Diff: Pre=0, Fol=3; Pre-Fol=0–3=–3 | | | | | |

MO = match-original: "correct" responses in all-capital letters.
fIa = interagreements with other responses in group.
SI = overall subject-interagreement: sum of fIa for group.
Like Sur = Pre or Fol responses paired with Sur responses.

Note in Table 1 that only three of 18 subjects guessed the original word, HIS, and all of them were in the Sur group: MO=3. In fIa scores, all Pre subjects received zeroes (they all guessed different words) but Sur and Fol subjects tied with SI scores of 8. And for Like Sur, Fol responses included three which paired with Sur responses (no response can appear in more than one pairing), but Pre subjects scored zero—as if, in Fragment #3, Fol context contributed more than Pre context to total Sur performance.

Sometimes subjects leave blanks vacant. For MO scores, of course, any non-response is "wrong." In scoring response interagreement, however, all non-responses are regarded as different not only from any given response, but also from each other. In this set—on, on, ___, ___, ___, ___—the first two responses would score fIa's of unity, but the last four would disagree with "on" and with one another. This rule also applies to Like Sur pairings; two non-responses do not make a pair.

## ANALYSIS AND RESULTS

### *Preliminary Note: Transformation of SI Scores to vSI's*

Unlike other quantitative measures in this report, SI scores do not yield raw data which can be entered directly into conventional analysis. Because subjects who interagree score duplicate fIa's, the only possible values of SI for a group of six subjects are 0, 2, 4, 6, 8, 12, 14, 20, and 30, and these nine values from an underlying scale of highly irregular intervals statistically difficult to deal with. However, the scale is at least monotonic (SI's increase without reversal from zero to maximum interagreement), so all SI scores were transformed into the ranks of their scale values (vSI's) from 1 through 9 before being used to assess correlations or differences in means. For example, SI's of 0, 2, and 30, respectively, transform into vSI's of 1, 2, and 9, which total 12 and have a mean of 4.

### *Preceding versus Following Contexts*

Table 2, in which the performance scores of all subjects and all context versions are summed for each of 45 fragments, shows no significant difference in any overall contrast of the means of comparable sets of Pre and Fol data—MO, vSI, and Like Sur. It seems fair to conclude that Pre and Fol contexts were about equal in their influences on cloze performance.

If some performance scores depend mostly on Pre context and others mostly on Fol context, it appears that there should be a negative correlation between them. Table 2 shows that all of the three Pre-versus-Fol correlation coefficients, *r*'s, are in fact negative, but none is individually significant. This suggests, however, that the projection of missing words may depend on cues both before and aft.

A closer (smaller-sample) look at Pre-versus-Fol match-original differences is presented in Table 3 for groups of only N=6 subjects. The groups administered Assemblies A and B scored higher on Pre than Fol, and for the latter the difference was significant. But Assembly C results were significant in the opposite direction, and Pre-versus-Fol differences washed out in an overall analysis of variance test.

### *Surrounding versus Preceding Following Contexts*

If the influences of Pre and Fol contexts actually are about the same, and each exhibits exactly half the amount of context presented by Sur versions, perhaps related Pre/Fol scores (based on the same fragments) summed over 12 subjects (two groups of six each) should tend to equal corresponding Sur scores summed over only six subjects. Would 12 responses wtih five words of context be equivalent to six

## Table 2

Comparisons, for $N=45$ Fragments, of Distributions of Scores for Paired Versions of Context*

| Versions of Context | Kinds of Scores | | |
|---|---|---|---|
| | Match-Original | vSI, Rank SI Values | Like Sur Pairings |
| Preceding: | 56 | 159 | 63 |
| Following: | 46 | 164 | 59 |
| Pre-Fol | 10 | –5 | 4 |
| Mean Diff | .22 | –.11 | .09 |
| SE (md)** | .43 | .56 | .39 |
| *t* w/44 df | .52 | .20 | .23 |
| *r* w/43 df | –.24 | –.12 | –.24 |
| Surrounding: | 121 | 237 | . . . |
| Pre/Fol: | 102 | 323 | . . . |
| Sur-(Pre/Fol) | 19 | –86 | . . . |
| Mean Diff | .42 | –1.91 | . . . |
| SE (md)** | .30 | .54 | . . . |
| *t* w/44 df | 1.41 | 3.53*** | . . . |
| *r* w/43 df | .59*** | .29 | . . . |

*Each distribution represented here combined scores from three groups, A, B, and C, whose assemblies administered different sets of 15 fragments for each context version.

**Standard error of the difference between correlated means.

***$p<.01$, 2-tailed; all other $t$'s and $r$'s have $p$'s$>.05$.

responses with 10 words of context as measured by MO and vSI scores? The answers are in Table 2.

For match-original scores, the Sur and Pre/Fol sums over the 45 fragments are 121 and 102, respectively, and their means do not differ significantly. However, the two distributions are positively and significantly correlated; $r = .59$. This positive correlation seems expectable with both Pre and Fol performances contributing to the Pre/Fol distribution of MO scores.

The comparisons of vSur and vPre/vFol distributions over the 45 fragments yield contrasting results. The correlation, $r = .29$, is not significant, but the vPre/vFol sum of 323 is significantly greater than the vSur sum of 237, $t=3.53$, despite the fact that interagreements on

Table 3

Match-Original Scores for Preceding and Following Versions of Context for Three Groups of Six Subjects Each

a. Comparisons Within Individual Groups

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A | 6 | 24 | 16 | 8 | 1.33 | .88 | 1.54 |
| B | 6 | 27 | 16 | 11 | 1.83 | .40 | 4.57* |
| C | 6 | 5 | 14 | –9 | –1.50 | .43 | 3.50** |
| Sum | 18 | 56 | 46 | 10 | | | |

*p<.01, 2-tailed
**p<.05, 2-tailed

b. Analysis of Variance for Repeated Measures for Three Groups of Randomly Assigned Subjects

| Sources of Variation | df | F |
|---|---|---|
| Between Assemblies (Groups), A, B, & C | 2 & 15 | 10.96* |
| Between Context Versions, Pre & Fol | 1 & 15 | 2.48 |
| Assemblies X Context Versions | 2 & 15 | 8.64* |

*p<.01, 2-tailed

"right" responses also contribute to SI and vSI scores. Perhaps the less constricting 5-word contexts somehow allowed subjects freedom to interagree on more "wrong" responses.

### *Match-Original versus Interagreement Scores*

Did subjects whose cloze responses agreed oftener with the author's original words also tend to interagree better with the responses of others?

Every subject's fIa's and corresponding "right" guesses were summed across his 15 responses to each of the three versions of context, and the computed correlation coefficients are entered in Table 4. All nine of the resulting *r*'s are positive and five of them are individually significant although each is based on only six pairs of scores. Then, using Fisher's *z*'-transformation, average coefficients were computed across all groups and across all versions of context, and all six of those *rz*' values were found to be significant.

The fIa sum for any individual subject has two components, one associated with "right" responses and the other with "wrong" ones. Of all 18 subjects, the best original-word guesser had an MO score of 18 out of 45. A count through the data showed that the subtotal of his fIa scores associated with MO responses came to 46, or 66.7% of his total fIa score of 69.

Similar counts were made for all subjects, and for every one the

Table 4

Correlation Coefficients: Match-Original versus Sums of *fIa* Scores for Individuals within groups

| Group of 6 Subjects | Context Version (Pairs of Responses Per Subject) | | | Average |
|---|---|---|---|---|
| | Pre (15) | Fol (15) | Sur (15) | $\bar{r}z'$ |
| A | .79 | .88• | .38 | .62• |
| B | .83• | .72 | .55 | .60• |
| C | .95•• | .92•• | .93•• | .89••• |
| $\bar{r}z'$ | .80•• | .78•• | .59• | |

•p<.05 (All probabilities are 2-tailed.)
••p<.01
•••p<.001

MO-related part of the total fIa score was the larger. The match-original percentages of fIa averaged over each group of subjects were 67.7% for A, 56.6% for B, and 64.3% for C. The overall average was 62.9%.

## DISCUSSION

This project failed to find any dependable overall difference between the influences of preceding and following contexts on subjects' guesses about the identities of words deleted from 45 prose fragments. Degrees of influence were scored by the number of subjects who successfully guessed the original words deleted and by the amount of interagreement among their guesses without regard for correctness.

The author has frequently been surprised by the surprise of others about his apparently unsophisticated lack of concern about possible differences in the influences of preceding and following contexts. However, although conventions about word order (from left to right in English and from top to bottom in Chinese) are important in the communication of information—"girl dream" is not the same as "dream girl" nor "dog house" the same as "house dog"—it seems obvious that one *thinks* both forward and backward. In "I heard a ______ . . .," how can a subject confidently guess "dog" or "bird" until he discovers that the blank is followed by "bark" or "chirp"? The writer is reminded that Shannon (1951) counted the number of trials it took a subject to correctly guess each of the successive letters in words and to spell words backwards. Subjects met with about equal success on the two tasks.

It hardly seems necessary to say that findings reported here should be considered tentative. They are based on a single set of prose fragments drawn systematically from a single rather colloquially phrased short story (Sherwood Anderson's "The Egg") and from the cloze responses, generated in about 20 minutes, of 18 subjects willing to spend half a class period playing a game of guessing specific words from what appeared to be wholly insufficient contexts. The author feels that these results should be generalized to English prose only to the extent thatt hey can be replicated with the use of other subjects, other stimulus materials, and other principal investigators.

## REFERENCES

Shannon, C. E. Prediction and entropy of printed english. *Bell System Technical Journal*, 1951, *30*, 50-64.

Taylor, W. L. Relative influences of preceding, following, and surrounding contexts on cloze performance. Paper presented at the meeting of the National Reading Conference, Tampa, Dec. 1971.

J. JAAP TUINMAN
Indiana University

# SPECULATIONS ON CLOZE AS A SEARCH PROCEDURE

Any production of words implies a retrieval of those words from memory. The selection of the appropriate elements is a function of contraints of varying degree. In free association these constraints are minimal; in responding to a question regarding one's age they are maximal. In a splendid paper, Wendell Weaver (1965) speculated about the cloze procedure in terms of the search and retrieval processes which must underly any production of a specific response to a involved in arriving at a correct response would amount to nothing deletion. As Weaver pointed out, a full description of the processes involved in arriving at a correct response would amount to nothing less than a ". . . rather complete psychology of language and thought involved in arriving at a correct response would amount to nothing [p. 127]."

It is one thing to say that clozing involves a search process; it is quite something else to specify the components of this process and their operations. Weaver in his description of such a search process did barely go beyond the observation that though the organism's searching is probably planned in some heuristic fashion, it is very difficult to start the study of this search plan from any other vantage point than the postulate of a logical exhaustive search.

What is obviously needed for a better understanding of the process characteristics of the cloze is a model of memory retrieval processes which has explanatory power for search activities presumably typical of cloze. In this paper I will attempt to show that an approximation to such a model exists. I will further go out on a limb by trying to analyze some of the implications of this model for our thinking about cloze as a process.

### A MODEL OF MEMORY SEARCH

The model of memory which I will briefly outline is described in detail by Atkinson and Shiffrin (1968a and 1968b), by Shiffrin (1968), and by Shiffrin and Atkinson (1969). Its essentials resemble those of the model proposed by Norman (1968).

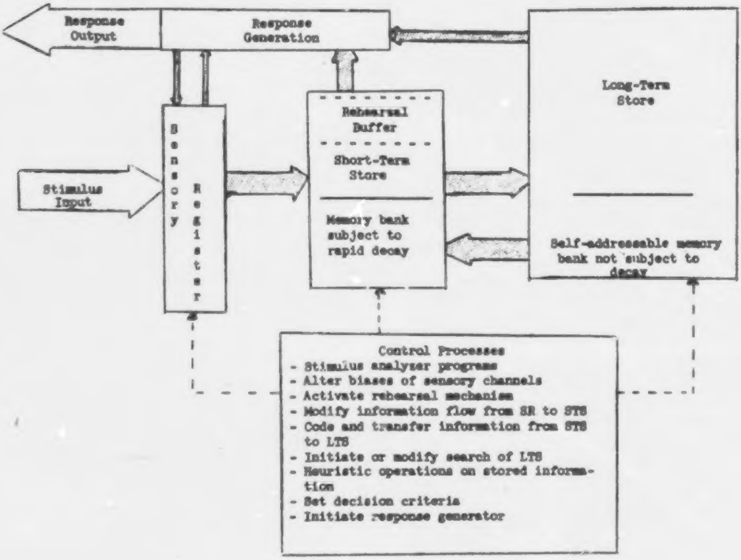

Response Output
Response Generation
Long-Term Store
Rehearsal Buffer
Short-Term Store
Memory bank subject to rapid decay
Stimulus Input
Sensory Register
Self-addressable memory bank not subject to decay
Control Processes
- Stimulus analyzer programs
- Alter biases of sensory channels
- Activate rehearsal mechanism
- Modify information flow from SR to STS
- Code and transfer information from STS to LTS
- Initiate or modify search of LTS
- Heuristic operations on stored information
- Set decision criteria
- Initiate response generator

Fig. 1. A flow chart of the memory system. Solid lines indicate paths of information transfer. Dashed lines indicate connections which permit comparison of information arrays residing in different parts of the system; they also indicate paths along which control signals may be sent which activate information transfer, rehearsal mechanisms, etc.

In the summary which follows I will discuss merely those character tics of the model which I consider particularly important in regard to cloze-search processes. Figure 1 depicts the model. In as far as the figure is not self-explanatory, the following comments may suffice. The sensory register holds incoming information for a very short time while the information is being transferred to the short-term store (STS). The STS will hold information no longer than 30 seconds unless the subject attends to it, for instance by initiating active rehearsal strategies. The STS is a working memory, transmitting to and receiving from the permanent long term memory stores (LTS).

In Shiffrin and Atkinson's view, the LTS is self addressing. That is, the "contents to be located themselves contain the information necessary to specify the storage location (s) [p. 181]." They compare such a system with a library where information retrieval is facilitated greatly by following the original storage plan: "An ensemble of information presented to the memory system will define a number of memory areas in which that information is likely to be stored; the memory search will therefore have certain natural starting points [p. 181]." The advantages of this model over location addressable or content- addressable systems which return respectively contents given certain locations or locations of given contents, are obvious. The degree of self-addressability, moreover, is variable: information presented to the LTS differs in the degree of storage specification.

In the Atkinson and Shiffrin model, retrieval is a three-component activity: search, recovery and response generation. The search process is recursive, it may be reinitiated after each successive response generation. Each cycle in the search process starts with the location of the next code (an ensemble of information that is closely related). This locating of codes (or images) may be either directed or random, depending on the search task. In directed search, a number of locations in some memory area are marked for examination. These locations are the examination subset. Various examination subsets may be examined subsequently. The recovery stage involves the pre-selection of information from the images located so that this information may be transferred to the STS. *Response generation* involves decision processes where the subject must decide whether or not the information retrieved is consistent with the test stimulus. If time permits, a subject may withhold a response and reinitiate search in the hope of finding a better response. Termination schemes may vary across subjects and across tasks. In general, decision processes differ for various tasks such as recognition and recall.

The test of any model is its capacity to take in data without internal bleeding resulting. Obviously my description of the Atkinson-Shiffrin model suffers from under-specification. However, when discussing the fit between some available data relevant to clozing as a search task and the model outlined above I will try to not violate the finer points of it.

The model calls for a memory file to be searched. This file is composed as a function of the (directed) search mechanism. For the moment we will assume that in the case of cloze such a file is made up of words, however stored. It seems furthermore reasonable to assume that with other things being equal both the success of the search and the time involved are a function of the size of the file. This must be true whether the search is self-terminating or whether it is exhaustive. Taylor (1954) noticed a high negative correlation between the probability of a correct response and the size of the corpus of words emitted as responses. In another experiment Taylor and associates (Taylor, Lore and Waldman, 1967; Taylor and Waldman, 1969) used unique and non-unique cloze items. Unique items permit only one response (____________ Fransisco) whereas non-unique items exercise less stringent constraints. By definition there is a difference in size of the number of target items in the search file. The prediction of shorter response latencies for unique items was confirmed for initial, unreinforced, testings.

From these two studies a number of things become clear. Cloze tasks differ from the paired-associate tasks which gave rise to the Atkinson-Shiffrin model in that the content of the examination subsets is not historically controlled by the researcher. The researcher must guess at the size and nature of the subset or search file for a given blank. He may, for instance, attempt to affect a dump of the file in a number of ways. He can either have a large number of subjects, supply one response to a given deletion or he can have few subjects emit many responses to that blank. In both cases the result is a distribution of words which, I assume, may be taken as an approximation of the file searched. The obvious shortcoming is that this distribution only contains items which have passed a criterion test in the response-generation phase. Yet, it is the closest we can come and inspection of response distributions does reveal many items which must be classified as nontarget, i.e. not fitting the constraints of the deletion in question. This suggests that in an aggregate of respondents some will apply criteria of fit in a loose enough way to let us examine non-fitting as well as target items.

A second observation flowing from the Taylor studies mentioned above is that one must constantly keep in mind that a cloze task not only requires the location of high-information images, the recovery of potential responses and a decision about whether or not a response fits the constraints. In addition, a typical cloze task demands that the S make a choice among possible alternatives so that he emits only the best response. I will return to this issue of a double decision process later.

In Atkinson and Shiffrin's view an examination subset resulting from the operation of the directed search mechanism will contain both target items (T) and non-target items (NT). As before, a target item is defined as one which fits the task constraints. From this view it can be predicted that an increase of non-target items in the file to be searched must be detrimental to both the speed of recovery of target items and the probability of a correct response. In cloze tasks the issue is complicated further by the fact that the probability of a correct response may be a function of the difficulty of deciding among near equally probable target items. I have some data which does show that the ratio of T to NT items is related to the probability of success but which does not discriminate among the two explanations just now advanced. In a recent study we analyzed the response distributions of 78 students to 51 noun deletions (Tuinman and Blanton, 1970). The data was gathered as part of a larger study described in Tuinman (1972). Responses were classified as T or NT on a judgment basis. Our intent was to relate the probability of a correct response (exact replacement) not only to the size of the corpus of emitted responses but as well to the T/NT ratio. From the response distributions the following statistics were calculated: k (= total word types); $k_1$ (= word types in NT-class); $k_2$ (= word types in the T class); $p_1 = k_1/k_2$; $p_2 = tk_1/tk_2$ where $tk_1$ refers to word tokens in word class $k_1$; p = proportion of correct responses. As in Taylor's study, the size of the distribution at each blank correlated highly with the probability of a correct response ($r = -.71$, $p > .01$). When not controlling for size of the distribution. the probability of a correct response was best predicted by equation 1 ($R = .75$)

$$p = .822 - .024k + .025p_1 \qquad 1)$$

A more interesting question in terms of a description of the examination subset, however, is which variable are powerful predictors of the response probability given a fixed size distribution. Deleting distribution size from the regression system resulted in equation 2

$$p = .773 - .014k_1 + .064p_2 - 1.566p_4 \qquad 2)$$

$p$ = proportion of subjects correctly filling in the blank
$k_1$ = size of distribution of NT types
$p_2$ = ratio of NT/T tokens
$p_4$ = ratio of T types and T tokens

The resulting multiple correlation is slightly lower ($R = .72$).

If indeed one is justified in treating the overt distributions as representative of the organization of the file(s) searched, the present results do support the contention that memory models derived from PA research may well be applicable to cloze research.

### THE DELETION AS A STIMULUS

I have, so far, avoided the problem of specifying to which extent encountering a blank and its linguistic environment is similar to instigators of search processes such as a single stimulus word in a PA task or the request to list all flowers one knows in restricted associative tasks. One might speculate that in a cloze test the task is less clear. more ambiguous and less constrained. This may be so in the case of some blanks, but it must generally be doubted that functionally, i.e. in terms of directing the search, an embedded deletion is such a unique type of search instigator that a general search model could not fit the cloze task. Weaver, Kingston and Dinnan (n.d.) had subjects emit responses to blanks for a duration of three minutes per item. After completing this task, each individual was presented with the same items but with one of his previous responses as a cue word. The cue was either the first or last word in a subjects' distribution. No significant differences were found between the total sum of the fluency scores under the no-cue and the cue condition. It must be noted in evaluating these results that the cued condition involved a repetition of the task as well. The cued condition used by these authors resembles the restricted associative response tasks originally used by Bousfield and his associates (Bousfield, 1944; Bousfield and Sedgewick, 1944; and Boufield. Sedgewick and Cohen, 1954). In these tasks the subject is required to list words in a specific category (flowers, cities, etc.). There has been a large amount of research on cumulative response curves relating the amount of words emitted to the rate of depletion of the total storage. This data regains importance in the light of the recent work on memory models. If it can be shown that cloze instigated search is but one of a number of special cases fitting a general model we will have gained much.

The finding of Weaver, Kingston and Dinnan intimates that at least at the level of production of alternate responses, pure cloze tasks and cued cloze tasks which resemble category constrained retrieval, result in distributions of similar size and general make up though differing in internal organization. The authors indeed do refer to vertical constraints on the response production as a function of the sematic category or categories from which words are retrieved.

Much stronger evidence that cloze (up to alternate response production but exclusive of the selection of a single response) closely matches category constrained search tasks is provided in a very important study by Kaplan, Carvellas and Metlay (1971). The authors had 9 to 12 Ss generate responses to blanks with context varying from zero up to 40 words. The purpose of the study was to relate rate and size of response production is a Bousfieldian equation of the basic form

$$F_t = T\ (1 - e^{-pt}),$$

Where F is thec umulative frequency of responses given by time t, T is the hypothetical upper limit of words recallable and p is a parameter which determines the proportional rate of depletion of the words. This basic equation which has been well researched with a number of different search tasks (such as category search; making up words from limited sets of letters) can be accounted for by a model of verbal recall such as presented by Shiffrin and Atkinson (1969). Its basic characteristics are the exponential form of the cumulative response curves and the inverse relation between rate of depletion and category size. Figure 2 summarizes the Kaplan and Carvellas' findings. The graphs speak for themselves.

With 40 words of context, T (total number of words retrievable) was only 19. With a decrease in context (10, 4, 2, 1, 0 words) T increased to 87, 149, 87, 213 and 746 words. Most important is that

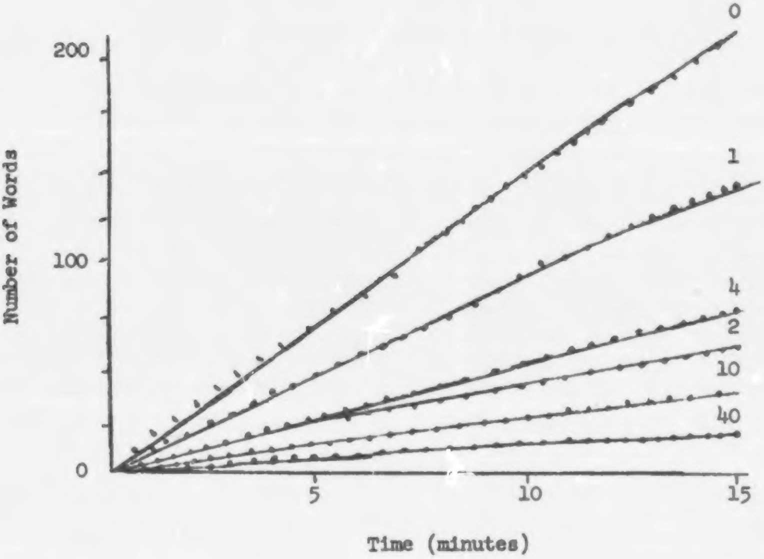


Fig. 2. Cumulative frequency of words produced as a function of time with 0-40 words of context. (From Kaplan and Carvellas, 1971)

though increase in context lead to a decrease in total words recallable together with an increase in proportional response rate, the change in (corpus recallable). This calls for a modification of the basic equation p (proportional response rate) was smaller than the change in T which assumes an increase of relative response rate proportional to the decrease in total words producible. This modification is necessary, according to Kaplan *et. al.*, because "Every word in the context presumably asserts some constraint on the choice to fill the blanks so that increasing context must increase the number of criteria that a potential response word has to satisfy. Therefore, the greater the amount of context, the more extensive testing of a trial word, and so the slower the scanning rates [1971, p. 211]." As they point out, the scanning rate should *not* change as a function of increased context if all a subject did was to check whether or not he had already emitted the response. On the other hand, the number of additional tests of the item against the constraints are a function of increased context. The necessity for these checks (demonstrated by a slower scanning rate) indicate that indeed the search file (the examination subset) must contain nontarget words as well as words fitting the constraints.

I find the Kaplan, *et. al.*, study very appealing. Application of the mathematical model underlying restrictive associative response tasks to cloze productions is not only attractive for research on the effect of linguistic variables on cloze responses but is an instrumental advance in terms of studying generation of cloze responses as a search proces. Their article is a demonstration of the latter ponit. The impact of their technique becomes even more clear in the light of the discussion of Indow and Togano (1970) in regard to the changes in values of the parameters of the Bousfield equation as a function of different assumptions about how items are retrieved from memory.

## RESPONSE GENERATION

An intriguing characteristic of a cloze task is that ostensibly a number of responses have similar privileges of occurrence in a blank. Is it necessary to postulate a two-phase decision process in the response generation phase? Are decisions made by first determining whether or not a retrieved item fits the constraints and subsequently by weighing the closeness of this fit?

Shiffrin and Atkinson (1969) indicate that the structure of the decision process varies with the task. As far as I know little is known about how Ss do proceed in regard to response generation in cloze tasks. If we assume that these decisions are made in short-term memory we may also assume the possibility of deliberate controll over these decisions on the subject's part. Recently Tuinman, Flanigan and Blanton in press conducted an experiment aimed at providing us with some information regarding the possible size of the set of alternative the subject was aware of choosing from. The subjects took regular cloze tests with modified instructions. Briefly, the task was to push a button everytime a word "occurred to the subject" as long as he did not settle on a response. The number of these reactions was recorded by an event-graph. A full description of the study is reported elsewhere but I will summarize the most salient points. As predicted, the number of button pushes was higher for difficult cloze passages than it was for easier ones. However, and this we found somewhat surprising, no blank solicited more than 5 responses. The majority of the blanks, around 75% overall, registered only *one* reaction. This is true, mind you, for a passage which was so difficult that only 6% of the answers were correct and also for a passage which had a cloze difficulty around 70%. Some of the passages had a preponderance of lexical deletions others contained mainly functional deletions. This too did not seem to make a difference.

You will notice that I reported the results in terms of button pushes, reactions, and the like, rather than in terms of words considered as possible responses. This does reflect a still existing hesitation on my part in regard to the validity of the task. Yet, both introspection and hard memory search data make our results quite plausible.

Take any cloze passage and try to complete it; as far as I can determine the response in most cases pops up automatically, very fast even in those cases where you're not sure it is the right response. In absence of a quick response a sort of emptiness filled by active rehearsal of part of the linguistic environment seems to exist. Only rarely do multiple responses seem to call for active and deliberate selection. So much for introspection. Atkinson and Shiffrin (1969) hold that during any search only a few images will be examined before termination. In an earlier report, (Shiffrin 1968) has estimated that Ss seldom consider more than five images from which to recover information. I am far from trying to better Miller and speculate about a magical number five. I am, however, a little more easy with our results than I would have been if a typical search in Shiffrin's paradigms had entailed examination sets of hundreds of items. I find the question of the size and the content of the set of items which are consciously considered by the respondent of interest for two reasons. Sternberg, Knoll and Nasto (1969 have shown that items in STS are processed at a slower rate than when in LTS because they are rehearsed. For the study of the cloze process that entails that one should attempt to break down response latencies into a STS and an LTS component. It seems reasonable to assume that the time in STS is, among others, a function of the number of alternatives considered. A second reason why I think it is important to determine to which extent the list of words occurring to you validly reflects a real component of the response generation process is the obvious access the researcher has to this list in terms of the elements contained in it.

## SUMMARY

To find out about process characteristics of the cloze we must think in terms of existing general models of information storage and retrieval. There is some data which indicates that a cloze task may indeed well be a special case which can be explained to a large extent by existing models. I explored that notion to some extent in reference to the Atkinson and Shiffrin model. In doing so I went out on a limb brushing over all kinds of questions of detail which must eventually be answered. I have tried to show that distributions of words fitting semantic and syntactic constraints as in cloze tasks may be a function of both horizontal and vertical constraints. Furthermore, the mathematical model describing the production of vertically constrained items can be modified to incorporate additional horizontal constraints as demonstrated by Kaplan and Carvellas.

## JUST A FEW OF THE QUESTIONS LEFT UNANSWERED

How are lexical items (content and function words) stored, so they are self-addressable? Is Osgood's notion of storage as sets of features helpful? (Osgood, 1968a & b; Smith, Osgood, Forster and Smith, 1971; Smith and Smith, 1971). What does an image or a code contain in terms of a cloze task? A number of sequentially related words? A word together with its semantic and syntactic feature descriptors? Are differences between pre-cloze and post-cloze a function of coding differences in term of temporal ordering of elements in the examination subsets? I share these questions with you merely to indicate how rich a field of cloze research lies yet untapped.

## REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: a proposed system and its controll processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivations advances in research and theory.* (Vol. 2) New York: Academic Press, 1968. (a)

Atkinson, R. C., & Shiffrin, R. M. Some speculations on storage and retrieval processes in long term memory. *Technical report 127,* Institute for Mathematical Studies in the Social Sciences. Stanford: Stanford University, 1968. (b)

Bousfield, W. A. An empirical study of the production of affectively toned items. *Journal of General Psychology*, 1944, *30*, 205-215.

Bousfield, W. A. & Sedgewich, C. H. An analysis of sequences of restricted associative responses. *Journal of General Psychology*, 1944, *30*, 149-165.

Bousfield, W. A., Sedgewick, C. H., & Cohen, B. H. Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology*, 1954, *67*, 111-118.

Feigenbaum, E. A. Information processing and memory. In *Proceedings of the fifth Berkeley symposium on methematical statistics and probability.* (Vol. 4) Berkeley: University of California Press, 1966.

Indow, T., & Togano, K. On retrieving sequence from long-term memory. *Psychological Review*, 1970, *77*, 317-331.

Kaplan, I. T., Carvellas, T., & Metlay, W. Effect of context on verbal recall. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 207-212.

Norman, D. A. Toward a theory of memory and attention. *Psychological Review*, 1968, *75*, 522-536.

Osgood, C. E. Interpersonal verbs and interpersonal behavior. *Technical Report 64*, Urbana: University of Illinois, 1968. (a)

Osgood, C. E. Toward a wedding of insufficiencies. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs: Prentice-Hall, 1968. (b)

Shiffrin, R. M. Search and retrieval processes in long term memory. *Technical Report 137*, Institute for Mathematical Studies in the Social Sciences. Stanford: Stanford University, 1968.

Shiffrin, R. M., & Atkinson, R. C. Storage and retrieval processes in long term memory. *Psychological Review*, 1969, *76*, 179-193.

Smith, S. W., Osgood, C. E., Forster, K. I., & Smith, W. R. Word retrieval as a function of semantic similarity of cue word. Paper presented at the meeting of the Western Psychological Association, April 1971.

Smith, W. R., & Smith, S. W. Effects of semantic feature similarity on false recognition. Paper presented at the meeting of the Western Psychological Association, April 1971.

Sternberg, S., Knoll, R. L., & Nasto, B. A. Retrieval from long term versus active memory. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1969.

Taylor, W. H. Application of cloze and entrophy measures to the study of contextual constraint in samples of ocntinuous prose. Unpublished Doctoral dissertation, University of Illinois, 1957.

Taylor, W. H., Lore, J. I., & Waldman, I. Relative latencies of semanticaphasics, stutterers, and normal controls for cloze items requiring unique and non-unique oral responses. Paper presented at the meeting of the Academy of Aphasics, Ann Arbor, October 1967.

Taylor, W. H., & Waldman, I. N. Latency and focus methods of cloze quantification. In G. B. Schick & M. M. May (Eds.), *Reading: process and pedagogy*. (Vol. 2) National Reading Conference 19th Yearbook. Milwaukee: the Conference, 1970. Pp. 241-262.

Tuinman, J. J. The removal of information procedure (RIP)—a first analysis. *Journal of Reading Behavior*, 1970, *2* (2), 44-50.

Tuinman, J. J., Flanigan, M. C., & Blanton, B. E. Subjects' awareness of response alternatives in cloze tasks. *Journal of General Psychology*, in press.

Weaver, W. W. Theoretical aspects of the cloze procedure. In E. L. Thurston & L. E. Hafner (Eds.), *The philosophical and sociological bases of reading*. National Reading Conference 14th Yearbook. Milwaukee: the Conference, 1965. Pp. 115-132.

Weaver, W. W., Kingston, A. J., & Dinnan, J. A. Vertical and horizontal constraints in the context of reading sentences. Unpublished paper, University of Georgia, n.d.

ERIC BROWN
New York University

# LEVELS OF SYNTACTIC REALIZATION IN ORAL READING

I would like to talk about two different perspectives in reading theory and research. The first I would characterize as the visual approach to the problem, where it is hypothesized that the fluent reader is able to go from the visual processing of letter configurations or features directly to some type of semantic interpretation, bypassing any need for auditory or articulatory mediation (Kolers, 1970; Smith, 1970). The second approach hypothesizes an additional step in this process, involving auditory rehearsal (Sperling, 1970), articulatory referencing (Hochberg, 1970), or abstract distinctive feature representations of lexical items (Chomsky, 1970; Brown, 1970). The former point of view is presently gaining much favor in professional reading groups, and before the issue is completely closed, I thought it might be profitable to review some of the evidence that bears on this question; especially since the reading texts and systems are already in press proclaiming this as the psycholinguistic theory of reading.

Let me say from the start that I know of no evidence that would suggest that it is possible to have a strictly visual representation of language, at any level. In fact the evidence is directly contrary. We do not know how we communicate by language, but it would appear that the auditory system is especially adapted to the rapid, sequential information-processing we call understanding language (Liberman, 1967; Lenneberg, 1967). The eye does not have the same temporal capacity for language as does the ear. Readout from very-short-term visual memory is relatively slow (Sperling, 1970), and even in visual-perception-to-writing tasks some type of auditory rehearsal component must be incorporated to account for what is retained. Unlike the earlier tachistoscopic studies, the study of sequential language displays suggests that visual masking and interference are factors in accounting for the slow, 3-4 item per second rate we find. We know, for example, that it is impossible to "read" any form of visible speech (such as spectogram) when held for continuous inspection, much less at real-time speeds.

We also know that in predicting letter reversal errors in early oral reading, differences in one phonetic feature are more important than problems in scanning. Liberman (1971) reports that in letter orientation confusions of the letters *b*, *d*, and *p*; *b* is most frequently mistaken for one of the other two consonants; perhaps because it allows the reader two chances to make an error by a single phonetic feature (place and voicing). It is also interesting to observe that there is no demonstrated relationship between visual perception and reading ability. For that matter, there is no evidence that one can train or improve visual perception itself, much less claim that it might be important in reading. On the contrary, some type of peripheral motor feedback or referencing appears to be more directly related to the central process.

The evidence on implicit speech or subvocalization in silent reading will not go away. It is embarrassing to all of us who favor a more cognitive view of reading, but I believe we are obligated to explain the phenomenon. The facts are that EMG recordings indicate tonic and phasic activity in the articulatory musculature during all silent reading (Edfeldt, 1960; McGuigan, 1970); that increasing reading difficulty is correlated with increases in this activity; that interfering with subvocal activity produces marked decrements in silent reading comprehension (Hardyck and Petrinovich, 1970); and that in general, subvocalization appears to facilitate reading comprehension. Various explanations have been offered for this correlated phenomenon, but all of them must account for the fact that articulatory referencing, even when no auditory signal is produced, facilitates the comprehension of language. The very interesting success of Rozin, Poritsky, and otsky (1971) in teaching Chinese ideographs to second graders with reading difficulties led them to propose not a whole-word or other visual approach to reading methodology, but instead, a return to Bloomfield's syllable as a suitable unit of instruction midway between the phonetic basis of English orthography and the morphemic character of the language. For despite what Smith (1971) and others of the visual processing persuasion have said about the lack of spelling-sound relationships in English, the orthography is not only essentially phonetic, but perhaps ideally so.

Chomsky and Halle in their major theoretical work on the phonological rules of English. *The Sound Pattern of English* (1968), have contended that an acceptable oral rendering of a passage requires greater linguistic sophistication, more fully involving the reading process, than was previously thought to be the case. They observe that "conventional orthography, being close to the linguistically significant system underlying ordinary speech, can be read only when the surface structure (including the internal structure of words) is known, that is, when the utterance is to some degree understood." This is again in counterdistinction to the visual processing hypothesis which contends that oral reading is irrelevant because it is possible to read aloud without comprehension, and that oral reading is grossly inefficient.

In an experiment reported more fully elsewhere (Brown and Miron, 1971), I have attempted to explore this hypothesis by establishing that a proficient oral reader demonstrates his grammatical complehension of a passage by the occurrence and duration of his pausing. The study specifically investigated the predictability of pause time in a 1537 word spoken message. The performance of a professional reader paced at 164 wpm was analyzed from three points of view. Three grammatical analyses of text were made. An IC, or Immediate Constituent, analysis was performed on each of the 84 sentences in the message. A simple IC boundary depth measure was calculated between each successive pair of words, counting all left and right facing brackets at that juncture. This measure would generally reflect surface structure complexity. A second measure (SCI) provided a slight variation on this procedure, following Chomsky and Miller's (1963) suggestion of a node-to-terminal-node ratio in the tree diagram of the terminal string. The third measure attempted to specify corresponding deep structure breaks in the surface structure. These deep structure analogues (DSA) generally coincided with clauses, however, additional specifications of other conjoining transformations were noted as well. This analysis was thought to account for deep structural breaks that might not occur in a surface structure analysis alone. An oscillographic recording of the entire message was performed, and text then appended so that all pauses could be related to morphemic analysis.

These three syntactic variables plus several additional lexical measures were entered into a multiple regression equation for the prediction of pause time. The final multiple *R* was .80, accounting for 64% of the pause time variance. IC accounted for more than 55% of the variance; DSA brought the $R^2$ figure to 61%; while SCI added an additional 2%. The important fact we draw from this display is that both IC and DSA are substantial predictors of Pause, and although they share a considerable proportion of variance (as would be expected from their theoretical assumptions) nonetheless each variable makes an additional unique contribution, preserving the grammatical distinction between observed and deep structure. It would appear that from a diagnostic point of view an acceptable oral reading performance indicated an understanding of both *deep* and surface structure configurations in the material read.

This level of understanding that is necessary for oral reading can be made quite precise. It does not mean that the message in its fullest sense is understood; there is too much evidence from oral reading problems to the contrary. However, it does mean that all grammatical relations are comprehended including the grammatical categories of words (Koler, 1970), and that the internal structure of the words themselves is well understood. (If we follow Chomsky, word meanings are undoubtedly a part of this understanding.) What is left is the psychological process of understanding language regardless of modality, a process undoubtedly made up of large components of context set, and attention. Nonetheless, it is at this level of processing that I have defined reading as a completed act.

The final issue I wish to raise is the problem of efficiency in reading. One of the major sources of evidence for the visual hypothesis is explaining the performance of rapid readers. Obviously someone reading at two to three thousand words per minute cannot be functioning at the speed of speech. Therefore they must be engaged in some more direct type of processing. Unfortunately, it is not clear that such reading is possible or has indeed ever taken place. We lack any systematic, well-controlled documentation of such performances, and the associated problems in measurement make prospects dim for such a study taking place in the near future. Moreover the work in time-compressed speech sugegsts that there is likely to be a uniform language processing speed, regardless of modality I am inclined to believe that our understanding of language and reading proceeds more or less at the rate our thinking about what we read or hear, which is the rate of implicit or explicit speech.

Certainly it seems premature to deemphasize the attention given to words and syllables in early reading, as is proposed in recent promotional literature for textbook concerns under the title of a psycholinguistic approach to reading. No one denies that semantic expectation and reduction of uncertainty are important strategies in successful reading. Neither is it unreasonable to encourage the largest manageable unit in the decoding process. But this assumes an automatization of function in reading that is not initially present. A skilled reader does most of his perception through the contextual constraints of language; but how does the child get there without principled reading acquisition instruction. To point to skilled readers as the model may be misleading. The fact that some children learn to read on their own without any systematic introduction to the process is interesting, but it is unrealistic to assume that all children might learn in this fashion. In Durkin's (1966) study of the spontaneous onset of early reading, she found that this ability was unrelated to intelligence within the normal to superior range and could occur in all social classes. Just as we used to have the phenomenon on radio of individuals who could do complex arithmetic operations in their heads—carrying out long multiplications without mnemonic aid—this does not mean that we should have all children aspiring after this ability using the same self-taught procedures.

In summary, I find no substantive evidence to suggest that a visual hypothesis in reading has any particular adequacy in explaining the available evidence. Surveying the research in visual information processing, oral reading, subvocalization, and reading speed, I find nothing to suggest that this theory is even commensurate with these findings. I hope that those involved with reading education will not converge prematurely on a theory of reading and reading instruction. I happen to believe that children need some systematic instruction in decoding print to oral language. To paraphrase Gattegno's recent assertion in the Harvard Educational Review, "The problem of reading is [not] solved."

## REFERENCES

Brown, E. The bases of reading acquisition. *Reading Research Quarterly*, 1970, *6*, 49-74.

Brown, E., & Miron, M. Lexical and syntactic predictors of the distribution of pause time in reading. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 109-118.

Chomsky, N. Phonology and reading. In H. Levin & J. P. Williams (Eds.) *Basic studies on reading*. New York: Basic Books, 1970. Pp. 3-18.

Chomsky, N., & Halle, M. *The sound pattern of English*. New York: Harper & Row, 1968.

Chomsky, N., & Miller, G. Finitary models of language users. In R. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical psychology*, Vol. II. New York: Wiley, 1965.

Durkin, D. *Children who read early*. New York: Teachers College Press, 1966.

Edfeldt, A. *Silent speech and silent reading*. Chicago: U. of Chicago Press, 1960.

Hardyck, C., & Petrinovich, L. Subvocal speech and comprehension level as a function of the difficulty level of reading material. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 647-652.

Hochberg, J. Components of literacy: Speculations and exploratory research. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970. Pp. 74-89.

Kolers, P. Three stages of reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970. Pp. 90-118.

Lenneberg, E. Prerequisites for language acquisition. *In Proceedings of the International Conference on Oral Education of the Deaf*, Vol. II. Washington: Alexander Graham Bell Assoc., 1967.

Liberman, A., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, *74*, 431-461.

Liberman, I., Shankweiler, D., Orlando, C., Harris, K. & Bell, F. Letter confusions and reversals of sequence in the beginning reader. Mimeographed paper, Haskins Laboratories, New Haven, Conn., 1971.

McGuigan, F. J. Covert oral behavior during the silent performance of language tasks. *Psychological Bulletin*, 1970, *74*, 309-326.

Rozin, P., Poritsky, S., & Sotsky, R. American children with reading problems can easily learn to read English represented by Chinese characters. *Science*, 1971, *171*, 1264-1267.

Smith, F. *Understanding reading*. New York: Holt, Rinehart, 1971.

Sperling, G. Short-term memory, long-term memory, and scanning in the processing of visual information. In F. Young & D. Linsley (Eds.), *Early experience and visual information processing in perceptual and reading disorders*. Washington: National Academy of Sciences, 1970.

ALBERT J. KINGSTON
WENDELL W. WEAVER
LESLIE E. FIGA
University of Georgia

# EXPERIMENTS IN CHILDREN'S PERCEPTIONS OF WORDS AND WORD BOUNDARIES

Teachers of reading, particularly teachers of beginning readers, tend to use the "word" as a major goal in reading pedagogy and for judging pupil progress. Constantly we see and hear references to word recognition, word attack skills, and sight vocabulary whenever reading is discussed. Interestingly little is known about how a beginning reader learns to recognize words or, for that matter, what a word actually is. It is difficult to reconcile the linguist's deliberate shunning of the concept "word", for he seemingly prefers terms such as phoneme and morpheme, witht he proclivity of the reading teacher and psychologist to use "word" as an analyzable language unit (Weaver, 1966; Weaver, 1967). At any rate "words" represent comparatively elusive units of study. particularly in oral language and even analytical philosophers have had difficulty in adequately defining the term. Because reading specialists apparently view words as so important to reading, the study of how children learn to distinguish words in both aural and written form should be helpful in understanding how children learn to read.

Karpova (1966) examined the ability of children aged three to seven to identify the number of words in a spoken sentence. The Ss were asked to tell the number of words they heard, and later to tell the ordinal position of words, i.e., which word was first, which word was second, etc. As a result of this study, Karpova hypothesized three stages of development: the youngest children used the sentences as "unified" wholes; older children were able to isolate the main action of the sentence upon repeated questioning by the examiner; a few older children were able to discriminate most of the words although they had difficulty with prepositions and conjunctions. Reid (1966) interviewed five girls and seven boys, five years of age, to discover what they though about reading. Among her findings was the fact that the five year olds showed a very poor understanding of the categories of "word" and "sound". Downing (1970) extended this study by presenting a non-human noise and four human utterances ranging from a single vowel phoneme to a complete simple sentence to Ss aged four to five years. Ss were instructed to say "yes" if what they heard was a word and "no" if what they heard was not a word. None of the children were able to identify a word as an adult would. McNinch (1970), as part of a broader study, constructed a so-called Aural Word Representation Test which required the Ss to duplicate the number of words heard by placing an appropriate number of foam rubber cubes equivalent to the number of words they thought they heard. He found a positive correlation between the AWR given in October and a standard end-of-year reading achievement test.

If one considers the philosophical studies of the concept "word" it is apparent that the treatment of the concept by linguists is not shared by most philosophical positions. Merleau-Ponty argues that the word has meaning, that a familiar thing appears indeterminant until we recall its name, and that the thinking subject is in a kind of ignorance of his thoughts until he has spoken or written them (Charlesworth, 1970). Among analytical philosophers the word as a meaningful unit has a high status (Hospers, 1967). Others argue that the word has a substantial character: one can hit a person with a word as well as with a weapon, and that at the beginning of conceptualization the word is considered an attribute of the thing itself (Blackmur, 1952; Burke, 1957; Werner and Kaplan, 1963).

The question of word boundaries had not received extensive examination from either philosophers or psychologists. Cowgill (1963) in a discussion of universals in diachronic morphology admits that in his study of word boundaries he was over-influenced by the printer's space of the text he used. He hypothesized that the identification of words is primarily a grammatical problem. Greenberg (1963) reached the same conclusion as a phonological problem.

In a study of children's discrimination of word boundaries in written language Meltzer and Herse (1969) sought to determine whether or not first grade children could locate the boundaries of written words in sentence form. They found that while most children identified some boundaries, there were many words the children could not identify using the criterion of the printer's space. The assumption here is that the identification of words is primarily a mechanical unitizing problem.

## THE EXPERIMENTS

In order to test the general hypothesis that first grade children do not understand the nature of a "word", a series of experiments were conducted which are briefly reported here.

### *Experiment One.*

This study replicated the Meltzer and Herse (1969) study but expanded the independent variable by using three distinct types of reading selections. Ss were 45 first grade pupils, upper-lower and lower middle class, randomly selected from 150 pupils enrolled in a rural, northeast Georgia school. There were three experimental conditions.

*The basal reader condition.* This consisted of 15 sentences or phrases selected from the Pre-primer and Primer editions of the basal reader series used at the school. The sentences or phrases ranged in length from a minimum of two words per sentence to a maximum of six words per sentence. Three sentences were presented at the beginning of the list in order to instruct the Ss, and these sentences were not scored or used in the final analysis. Each sentence was typed, using a Primary typewriter, on 1" by 8" strips of paper.

*The second-order approximation to English condition.* Nonsense words were formed using the basal reader material as a source. A letter was located in a random position on a page. This letter was then located at its next appearance in the text and the letter next to it recorded. This letter was then searched for next in the text and the letter next to it recorded, Constraint was imposed upon this approximation by fixing the length of the pseudo-words equal to the length of the words in the first experimental condition. These pseudo-sentences or phrases were also typed on individual strips of paper.

*The adult novel condition.* Sentences were selected from an adult novel. The number of words per sentence was held constant to match the number of words per sentence in the other two experimental conditions. The number of letters per word was greatly increased. These sentences were then typed on strips of paper.

The subjects were obtained from their classrooms individually, given a pair of scissors, and asked to cut the strips of paper where ever they thought a word occured. There were 15 Ss in each experimental group. The first three sentences administered to each Ss were used to illustrate and to make sure the Ss understood and could perform the cutting task. The order of sentences in each condition was randomized for each Ss.

In analyzing the results the scoring criteria developed by Meltzer and Herse (1969) were adopted. These were:

Type A error–divided word at tall letter.
Type B error–divided word elsewhere than at tall letter.
Type C error–combined two words.
Type D error–combined letters with no regard to space.

Type E error—equated words and letters.
Type F error—other combinations.

In analyzing the responses it was noted that a number of errors did not fit the criteria. Such errors were placed in the other combinations category.

For each Ss the different types of errors as well as the frequency of each type of error was recorded. (Tables available from authors upon request.) The total number of errors for all Ss was then calculated and means, variances, and standard deviations computed (see Table 1).

Table 1

Means, Variances, and Standard Deviations of Errors for Various Materials in Experiment One

| Materials | Means | Variance | Standard Deviation |
|---|---|---|---|
| Basal Reader | 5.27 | 58.21 | 7.63 |
| Second-Order Approximation to English | 6.53 | 84.14 | 9.17 |
| Adult Novel | 20.80 | 991.71 | 31.49 |

A significant difference was found between the Adult Novel task and both the Basal Reader and the Second-Order Approximation tasks. An F-value of 5.64 was attained.

Table 2 shows the errors attained using the Meltzer-Herse system of classification. It can be noted that a number of different types of errors were found.

Table 2

Total Frequency of Errors Over All Three Conditions
(Values indicate number of times a particular type of error occured)

| Conditions | No Errors | Divided at Tall Letter | Divided Elsewhere Than at Tall Letter | Combined Two Words | Combined Letters With No Regard to Space | Equated Words And Letters | Other Types |
|---|---|---|---|---|---|---|---|
| Basal Reader | 4 | 5 | 5 | 43 | 6 | 2 | 3 |
| Second-Order Approximation to English | 2 | 6 | 12 | 52 | 12 | 0 | 1 |
| Adult Novel | 1 | 80 | 75 | 84 | 47 | 2 | 9 |

The most common error was that of combining two or more words. For the basal reader and second-order approximation conditions this was the major source of error and the combining errors were generally restricted to combining two words of which one word was a one-letter word, for example; "I said", "I can," "and I." On the other hand there were some instances when combinations such as; "said, 'No' ", "get the", and "is not", were observed. This type of error carried over to the adult novel material condition where a number of other word combinations were introduced, for example; "lowered his", "her foot impatiently", "what you", "his knee", and "even more pronounced". The combinations here seem to be a result of a failure to perceive any word meaning, in addition to the fact that the printer's space was not recognized as a word boundary cue. The adult novel material brought out high frequencies of all the Meltzer and Herse categories except the "equating the letter and word" category and the "other category". There was great confusion about word boundaries; some of them were divided at ascenders and some at descenders, and some at neutrals. There also was the error of putting together parts of two words, i.e., the last letters of one word and the first few letters of the next word, for example; "mo/re pr/onoun/ced".

It should also be observed that there is no difference in the response to basal reader materials which were constructed from materials which the Ss covered in their classrooms, and a second-order approximation to English equated to the word length of th e basal reader materials. The error patterns in the second-order approximation materials is essentially the same as that of the basal reader materials with the preponderant error being a combination of two words, in most cases one of the words consisting of a one letter word. for example; "vex i", "iro s", "i ner", and "c apro."

It would appear from the results of this experiment that recognizing the printer's space as the seperator of words is secondary to perceiving that a particular linguistic unit represents a meaningful entity. The reason for the performance in the case of long words is obscured because length of the word and the ability to perceive words as meaningful units are correlated—the longer the word the more difficult is its comprehension. In this case therefore there may be a regression to a more primitive level of perception and analogies drawn to the basal reader situation where words of six and seven letters in length are rare. This "stimulus-bound" perception would then also account for the performance on the second-order approximation materials. Here the length of the pseudo-words are the same as that of the basal reader materials. And the behaviors toward the two kinds of materials are essentially the same. Thus, the developmental trend observed by Meltzer and Herse (1969) was not found in our data.

*Experiment Two.*

In experiment two, the materials used were taken from a aural word representation task presented originally by McNinch (1970). The stimulus consisted of three short trial sentences and sentences or phrases at each length from two to six words per sentence or phrase.

There were three experimental conditions. The Ss heard the examiner read the sentences to him; the Ss were presented with the sentences typed with a primary typewriter on strips of papr; and the Ss heard the sentences read to them via a tape recording. Each Ss was given a number of wooden cubes and was instructed when he heard or read a sentence to place the number of cubes equivalent to the number of words in the sentences before the E. The dependent variabie was the number of cubes successfully placed.

As shown in Table 3, the highest scores were obtained over the visual presentation. This "reading" condition also had the most constricted spread of scores as shown by the standard deviation. (Tables illustrating these data showing analyses by frequency of errors available upon request.)

Table 3

Means, Variances, and Standard Deviations for Conditions One, Two, and Three in Experiment Two

| Conditions | Mean | Variance | Standard Deviation |
|---|---|---|---|
| Aural | 7.38 | 25.27 | 5.03 |
| Visual | 12.07 | 6.87 | 2.62 |
| Taped | 6.20 | 12.40 | 3.52 |

A simple, one-way, randomized design, analysis of variance was computed. The F-value was significant beyond the .05 level of confidence.

Table 4 is a summary of successes and errors of all Ss in judging the number of words in a sentence or phrase. The most important consideration is the columns labeled Times Below and Times Above the correct number of words in the stimulus. In the reading aloud and taped conditions the number of words were consistently underestimated while in the reading condition, which had many fewer errors, the number of words in the stimulus were overestimated.

It seems that for these young Ss that amalgamation of units is characteristic. Ss however reading the same materials seem to have many more cues to the word unit as revealed by their greater success in counting the correct number of words and when they commit an

Table 4
Summary of Successes and Errors in Sentence of All Subjects in Experiment Two

| Conditions | Number Correct | Number Incorrect | Times Below | Times Above |
|---|---|---|---|---|
| Aural | 95 | 130 | 119 | 11 |
| Visual | 166 | 59 | 16 | 43 |
| Taped | 78 | 147 | 115 | 32 |

error the tendency is to perceive more units then those defined by the printer's space rather than less.

*Experiments Three, Four, and Five.*

The results of experiments one and two as well as the work of Reid (1966) and Downing (1970) and others led us to suspect that one source of difficulty that children encountered in recognizing word boundaries was due to their faulty concepts concerning the nature of a "word." Careful study of philosophical, linguistic, and reading literature reveals that the construct "word" is a vague term. Unfortunately time and space limitations preclude an adequate discussion of just how ambiguious the term is.

Experiment three tested the children's ability to identify human words and human sounds. A tape consisting of common words and common human utterances was prepared. A group of 15 randomly selected first graders were asked to identify the sound they heard as either a word or not a word. Experiment four consisted of 10 compound words and 10 two word pairs, presented on a tape to another group of 15 first graders. The Ss were asked to identify what they heard as either one word or two words. Some of the "sounds" used in experiment three were: hum, ahem, mmm; tsk; oh. Some of the words used were map; amen; shoot; street; banquet. For experiment four some of the compound words were policeman; skyscraper; baseball; quartermaster. Some of the two word pairs were: ball game; two wheeler; dump truck; fire engine.

As shown in Table 5, the children as a group had some difficulty in telling the difference between human words and human utterances and in recognizing compound words from two words.

The fifth experiment consisted of twelve taped sentences which contained both sounds and words. Fifteen first graders randomly selected were asked individually to tell how many words they heard in each sentence. Examples of the task were: (Snore) (Snore) Daddy is asleep; Listen to the teapot. (Whistle); (Arf) (Arf) went the dog.

Both human and inanimate sounds were used and the position of

Table 5

Means, Variances, and Standard Deviations for the Word-Sound Test and the Compound Word Test

| Test | Mean | Variance | Standard Deviation |
|---|---|---|---|
| Words | 6.67 | 4.29 | 2.07 |
| Sounds | 7.33 | 1.00 | 1.00 |
| Total (Words & Sounds) | 14.00 | 5.00 | 2.24 |
| Compound Words | 5.40 | 8.86 | 2.98 |
| Two Word Pairs | 6.73 | 7.36 | 2.71 |
| Total (Compound & Two Word Pairs) | 12.13 | 10.86 | 3.30 |

the sound in relationship to the sentence was varied. Unfortunately this task was too difficult for these Ss. It seemed apparent however that many chldren had difficulty in not only remembering what they heard and counting the number of words, but also confounded noises and sounds with the words in context.

## CONCLUSION

These five experiments demonstrate quite conclusively that first grade children lack precise concepts concerning the nature of a "word". Apparently in oral language the concept is not a major barrier either to speaking or listening. In reading however word recognition is an essential and fundamental skill. Perhaps the difficulties of certain children in learning to read may stem from the problem they have in distinguishing words. It also is possible that young children may attend to word units when they begin instruction in reading. Until that time there apparently is little need to concentrate on separating words from their context. As noted in experiment one, however, this does not mean that learning the printer's space is the primary requisite for unitizing words. After all, the concentration in reading instruction is not upon identifying word boundaries, but upon perception of words themselves. It is also impressive that once children learn the conventions for identifying word units in reading (an ability which comes with the mastery of reading skills) it is apparently an easy transfer to break up the aural phonological stream into word units. These first graders however, do not exhibit this easy transfer which is apparent in older children. It is possible then that this is because they are still erratic in their perception of both meaning and the function of the printer's space, and therefore do not have a firmly grounded reading perception of words from which to operate on the phonological stream.

One final speculation seems to be in order, although these experiments did not address themselves to this question. As one observes children grappling with the types of tasks presented here, one can only wonder about the nature of the "word" as a unit of language. Linguists who as a group prefer to work with oral language tend to shun the construct "word." These studies tend to show that for young children at least the word does not seem important. Perhaps sounds become words only in written form, and perhaps even in written form they are not too important.

## REFERENCES

Blackmur, R. P. *Language as Gesture*. New York: Harcourt, Brace, and World, 1952.

Burke, K. *The Philosophy of Literary Form*. New York: Vintage Books, 1957.

Charlesworth, J. H. Reflections on Merleau-Ponty's phenomenological description of "word". *Journal of Philosophical Phenomenological Research*, 1970, *30*, 609-613.

Cowgill, W. A search for universals in Indo-European diachronic morphology. In J. H. Greenberg (Ed.), *Universals of Language*. (2nd ed.) Cambridge: The MIT Press, 1963.

Downing, J. Children's concepts of language in learning to read. *Educational Research*, 1970, *12*, 106-112.

Greenberg, J. H. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of Language*. (2nd. ed.) Cambridge: The MIT Press, 1963.

Hospers, J. *An Introduction to Philosophical Analysis*. Englewood Cliffs: Prentice-Hall, 1967.

Karpova, S. N. Osoznanie slovesnogo sostava rechi rebenkom doshkol'nogo vozrasta (The preschooler's realization of the lexical structure of speech). Voprosy. Psikhol. 1955, *4*, 43-55. Abstracted by D. I. Slobin, In F. Smith & G. Miller (Eds.), *Genesis of Language*. Cambridge: The MIT Press, 1966. Pp. 370-371.

McNinch, G. H. W. The relationship between selected perceptual factors and measured first grade reading achievement. (Doctoral dissertation, the University of Georgia) Ann Arbor. University Microfilms, Order No. 71-3756. *Dissertation Abstracts*, Feb. 1971, *31*, 3965A.

Meltzer, N. S. and Herse, R. The boundaries of written words as seen by first graders. *Journal of Reading Behavior*, 1969, *1*, 3-13.

Reid, J. F. Learning to think about reading. *Educational Research*, 1966, *9*, 56-62.

Weaver, W. W. Units of measurement and units of language. Paper presented at the symposium on *The technology of written instruction*. New York: Columbia University, 1966.

Weaver, W. W. The word as a unit of language. *Journal of Reading*, 1967, *10*, 262-268.

Werner, H., & Kaplan, B. *Symbol Formation*. New York: John Wiley and Sons, 1963.

GEORGE McNINCH
University of Southern Mississippi

# PARADIGMATIC LANGUAGE TRAINING AS A METHOD OF INCREASING READING ACHEVEMENT

Language factors play a significant role in activities which encompass reading tasks. It is impossible to consider reading independent of language features or patterns. "Reading is a process of recognizing that printed words represent spoken words and is a part of the total language spectrum [Gunderson, 1970, p. iii]."

Recently, Dinnan (1971) has postulated that specific language behaviors are very closely tied to reading. In fact, Dinnan has theorized that one language category, paradigmatic responding (responding with a word in the same form or class as the stimuli word) may be the key to learning. The language skills that control a child's ability to function in reading are the skills that allow him to make convergent (paradigmatic) and divergent (syntagmatic) language responses to an instructional or reading situation. Language manipulations, organized response selections or mediations, involving paradigmatic operations allow the child to communicate and think in the same language mode as the instructor, thus fostering a successful educational interchange. Syntagmatic responses, unorganized, parallel language responses, possily interfere with the learning set. Divergent reaction patterns are established thus the communication link is jeopardized and content is not fully expanded or exploited.

## BACKGROUND

A substantial data base is being collected that demonstrates the close relationships between paradigmatic responses and reading achievement or paradigmatic language and intelligence. Several studies using word association measures (McNeil, 1966; Ervin, 1961; Entwisle, Forsyth, & Muuss, 1964; and McNinch, Richmond & Hornsby, 1970) have outlined the close relationship between the frequency of paradigmatic responses and age progression. Paradigmatic responding may be a developmental task controlling the pupil's ability to master material presented at age or grade levels.

Specific word association measures, e.g., responding paradigmatically or syntagmatically to stimuli words, have been significantly related to achievement and ability measures. Bickley, Dinnan and Jones (1971) determined that first grade pupils who responded syntagmatically to a word list did significantly less well than paradigmatically responding pupils on a standardized readiness test. R. Bickley, Bickley and Cowant (1971) working with fourth grade pupils hypothesized and confirmed that pupils who responded paradigmatically were better reading test takers than pupils who responded syntagmatically. Using an eighth grade sample, Shen and McNinch (1971) found a significant difference in the frequency of paradigmatic responses between good and poor reading groups. Significant differences were also noted when comparisons by race, but not by sex, were made. Dinnan, Bickley and Bickley (1971) found that more mature subjects, 19 year olds, also responded paradigmatically or syntagmatically in relation with their reading achievement. Dinnan, Cowart, and Bickley (1971), and Dinnan, Bickley and Williams (1971) found paradigmatic responding to be closely associated with economic status and intelligence, respectively. After summarizing available literature treating language responses Dinnan (1971) concluded:

> When a measure of success was placed against the type of responses, a clear (.01 level of confidence) indication was that those subjects who produce a 'higher level" (superordinate, coordinate, whole-part, or contrast) product, as measured by the "products of the intellect" Guilford (1967) were making a success in formal scholing. Those subjects who did not make this type of response were far below any form of success (p. 3).

The ability to produce paradigmatic language responses clearly seems to be linked to measures of achievement, ability, or other forms of success. However, several basic questions concerning the importance of this word association measure still remain unanswered: Is paradigmatic production a basic skill separable from measures of basic ability or academic sucess? Is paradigmatic language training the therapy necessary for building the intelectual operations (products) and comunication tools requisite to the promotion of reading skill development in groups of disabled readers?

## THE PROBLEM

Though the relationships betwen word associations and achievement has ben amply demonstrated, no specific study has yet sought to isolate the existence of paradigmatic responding as a skill area separate from academic or basic intelectual skills. As yet, no study has been conducted to determine the effect that training in paradigmatic language responses would have on reading achievement. An indepth study of this type would yield specific data on the nature of paradigmatic responses, and it could develop the relationships existing between the quality of language responses and the skills evaluated through measured reading achievement.

The basic questions posited for this study were used to generate the following hypotheses:

1) There are no significant relationships between paradigmatic language responding and measured skills of reading or intelligence.

2) Training in paradigmatic language responses does not significantly effect achievement in word meaning.

3) Training in paradigmatic language responses does not significantly effect achievement in paragraph meaning.

4) Training in paradigmatic responding does not significantly effect measured achievement in paradigmatic responding.

## PROCEDURES

### *Sample and Pretesting*

The subjects included in the study were 45 sixth grade pupils reading below grade level from a Mississippi Junior High School. The 45 pupils were randomly selected from a sixth grade population of 180 pupils observing the criteria that they be one to thre years retarded in reading as measured by the *Stanford Achievement Test (Reading) II, Form W;* responding with more than 50% synytagmatic responses to the *Oral C/A Language Inventory;* and within the range of usual classroom intelligence (80-120) as measured by the *Lorge-Thorndike Intelligence Test, Level D.* Stimulus words used on this Oral C/A Language Inventory were:

1. book
2. king
3. joy
4. arm
5. life
6. soul
7. pleasure
8. river
9. fact
10. apple
11. basement
12. anger
13. earth
14. belief
15. door
16. chance
17. wife
18. opinion
19. lip
20. deceit
21. effort
22. sky
23. gaiety
24. tree
25. idea
26. woman
27. truth
28. cash
29. gift
30. Happiness

Pretesting of the parent population began in September, 1971, and lasted one week. The *Lorge-Thorndike* and *Stanford* instruments were given according to the instructions in their respective manuals. Guidelines concerning group size, length of testing period, and general testing coditions were closely followed. Testing on the language inventory was done in group sessions after the training groups had been formed.

The sample included 25 boys and 25 girls. Mean I. Q. for the group was 92 with a mean chronological age of 11.4 years. The group had a mean of 23 syntagmatic responses on the 30 item language measure.

*Methods and Training*

Following the population pretesting, the selected 45 subjects were randomly assigned to one of the three treatment groups: experimental; placebo; and control.

The experimental treatment developed for the investigation was composed of 12 structured lessons based upon Dinan's (1971) levels of paradigmatic language performance.

Table 1

Sample Experimental Lessons

| | |
|---|---|
| **In a 10 minute period, given 20 pairs of words, each student will write the relationships of at least 5 pairs of words.** | **Distribute mimeographed copies of worksheet containing 20 pairs of equals. Place two examples on chalkboard such as:**<br>**1. cat – tiger**<br>**2. car – train**<br>**Discuss existing relationships. Allow time for written. work. Discuss individual answers and different relationship possibilities.** |
| **In a 10 minute given period, given 10 pairs of words each student will match five pairs of related words in writing.** | **Distribute mimeographed copies of worksheet containing 20 pairs of equals.**<br>**Place two examples on chalkboard, such as:**<br>**1. cat bicycle**<br>**2. car dog**<br>**Discuss matching procedures. Allow time for written work.** |
| **In a ten minute period, each student will write at least three sentences using each of a pair of opposites in each sentence.** | **Motivation: Ask "What is the opposite of grimy?" Accept any reasonable answer such as neat. Ask two children to use both words in one sentence.**<br>**Example: "The grimy boy looked extremely neat when Mrs. Jones finished with him."**<br>**Distribute mimeographed sheets containing ten word contrasts:**<br>**parent – child**<br>**noisy – quiet**<br>**bright – dull**<br>**truth – lie**<br>**forward – reverse**<br>**shakey – calm**<br>**gigantic – small**<br>**deep – shallow**<br>**love – hate**<br>**now – later**<br>**Ask students to make one sentence using each pair of contrasts.**<br>**Individual students will read a sentence that they have composed. Let class discuss different possibilities for each pair.** |

Table 2

Sample Placebo Lessons

| | |
|---|---|
| To verbally contribute at least one sentence to a story built upon by each member of the class during a 15 minute period. | The teacher explains to the class that they will make up a story. Each member will contribute. The teacher or one of the students begins the story. Each student adds at least one sentence, verbally to the story. Students decide on a title. |
| Each subject will write on paper the mental image elicited by fifteen sounds presented on tape in 20 second intervals. | Teacher will pre-record on tape eighteen different sounds. Example: ice clicking in glass, running water, and whirling eggbeater.<br>Play one sound. Ask children to orally give mental images produced. Allow 4-6 students to respond. Repeat procedure three times.<br>Play sounds one at a time at 20 seconds intervals. Children will write on paper the mental image which the sound stimuli evokes in them.<br>Repeat sounds one at a time, allowing different child to read their written responses to it. |
| Each student will write on paper two rhyming words to be used as a stink-pink" within a fifteen minutep eriod. | Explain to the class the rhyming pattern of words which sound alike and have the samen umber of syllables.<br>Give them examples, as stink-pink, red-bed.<br>Tell the students to think of a pair of words and a definition to explain it. Example: Something made of a metal that you sleep on. Answer: lead-bed.<br>Have each student write two pairs of rhyming words on paper. Each student will present one definition for one pair to the class. The class will have one minute to guess the "stink-pink." |

Each of the 12 experimental lessons lasted approximately 30 minutes at the same time each day (8:10 to 8:40). The four levels of paradigmatic functioning were treated by three structured lessons each. The lessons incorporated written and verbal development of paradigmatic language. Sample lesson plans can be found in table 1. The placebo group received 12 thirty minute lessons of a language enrichment nature (but not paradigmatically oriented) distributed over the same three week period. Table 2 displays sample placebo lessons. The control group received no instruction, only pre- and posttesting. All treatments were administered by two female graduate students in reading. The two rotated between the experimental group and placebo group in order to minimize possible teacher effects. The experimental and placebo training was supplementary to regular classroom reading instruction.

### *Posttesting and Analysis*

Posttesting was conducted at the end of the treatment period. Posttesting instruments were an alternate form of the Stanford, Form X, and a readministration of the language inventory. Testing was done in Testing was done in group sessions by the treatment instructors. Posttesting conditions were not ideal; noise and movement from oth[illegible] students may have distracted the subjects to some extent. Specific dire[illegible]tions were followed in test administration.

Intercorrelation analysis was the statistical tool chosen to evaluate the first hypothesis. Analysis of covariance was used to test the second, third and fourth hypotheses formulated for the study. The pretest reading scores and the pretest language score were used as covariants in their separate analyses. Covariance was selected as the statistical tool to test the hypotheses since it would allow for control of pretest experience on learning.

## RESULTS

The first investigation attempted to establish the separatness of paradigmatic responding as a skill unique from academic or intellectual features. The computed correlations, depicted in Table 3, indicated that paradigmatic responding was nonsignificantly related to the academic (word meaning and paragraph meaning) or intellectual variables. In fact, word meaning was negatively correlated (–.23) with paradigmatic language and the correlation between the language measure and paragraph meaning was a slight .19. The highest correlation, .37, was found between the language measure and intelligence. To clarify the nature of the correlations, coefficients of determination were computed. The $R^2$s indicated that the percentage of related variance among the variable (Table 4) was quite low. Therefore, the null hypothesis (There are no significant relationships between paradigmatic language responding and measured skills of reading and intelligence.) was not rejected. Since the associated percentages of commo[illegible]riance were quite low, it may be tenable to conclude that paradig[illegible] responding is possibly a unique skill, related to, but no synonym[illegible] with academic or intellectual ability.

Table 3
Intecorrelations Among the Variables of Paradigmatic Responding, Reading Achievement, and Intelligence

| Source | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|---|---|---|---|---|
| $X_1$ Paradigmatic Responding | – | –.23 | .19 | .37 |
| $X_2$ Word Meaning | | – | .17 | .05 |
| $X_3$ Paragraph Meaning | | | – | .32 |
| $X_4$ Intelligence | | | | – |

The null hypotheses of no difference on reading achievement between groups after treatment were tested using word meaning and

Table 4

Coefficients of Determination Among the Variables of Paradigmatic Responding, Reading Achievement, and Intelligence

| Source | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|---|---|---|---|---|
| $X_1$ Paradigmatic Responding | – | .052 | .036 | .144 |
| $X_2$ Word Meaning | | – | .028 | .002 |
| $X_3$ Paragraph Meaning | | | – | .102 |
| $X_4$ Intelligence | | | | – |

paragraph meaning as the criterion variables. Comparisons were

paragraph meaning as the criterion variables. Comparisons were made across control, placebo, and experimental groups. In the analyses, pretest scores were used as the respective covariants. As indicated by the covariance analysis in Table 5, there were no significant differences between the groups on the criterion variable word meaning; therefore, the second null hypothesis was not rejected. There were no significant differences whene paragraph meaning was the criterion variable in the covariance analysis. Therefore, the third null hypothesis was also accepted. Mean scores across the groups were not different. Generally it can be concluded that the generated language training for this select sample did not facilitate academic (reading) success. Neither the experimental treatment nor the enrichment placebo exercises seemed to promote significant growth in reading achievement over the control group.

Table 5

Analysis of Covariance

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Word Meaning | | | | |
| Treatments | 2 | 87.80 | 43.90 | 1.09 |
| Within | 39 | 1564.58 | 40.11 | |
| Paragraph Meaning | | | | |
| Treatments | 2 | 170.78 | 85.39 | 1.45 |
| Within | 39 | 2293.79 | 58.81 | |
| Paradigmatic Responding | | | | |
| Treatments | 2 | 128.63 | 64.32 | 14.62* |
| Within | 39 | 171.54 | 4.40 | |

* $p < .01$

The fourth hypothesis on paradigmatic responding was also tested by analysis of covariance. A significant F was recorded. The groups did seem to be responding differently to the paradigmatic inventory after treatment. To determine the source of this significant variation, Scheffe's test was used to make simple comparisons on the adjusted mean scores. Table 6 indicates that the sources of significance variation were distributed among all three groups. Significant F scores beyond the .01 level of probability were noted in each simple comparison. Apparently each group was responding in a quantatively different manner in regard to mean numbers of paradigmatic scores given. Observation of the adjusted mean scores in Table 6 indicate that after treatment the experimental group had the highest mean score, and the placebo group had the lowest mean paradigmatic score. The control group, which had received no training, was higher than the placebo. The null hypothesis of no difference aomng groups on paradigmatic responding after language treatment was not accepted. The groups did differ. It may be tenable to conclude that paradigmatic responding is a skill that can be taught if training (experimental) is directly guided to this end. Moreover, mere language enrichment exercises (placebo) do not seem to foster growth in this specific language area.

Table 6

Scheffe Comparisons (Adjusted Means) on Paradigmatic Responding

| Comparison by groups | Difference | F | p |
|---|---|---|---|
| Control vs. Placebo | 2.076 | 7.101 | .01 |
| Control vs. Experimental | –2.141 | 7.299 | .01 |
| Placebo vs. Experimental | –4.271 | 29.30 | .01 |

| | Control | Placebo | Experimental |
|---|---|---|---|
| Adjusted Means | 7.818 | 5.742 | 9.959 |
| Unadjusted Means | 7.286 | 6.067 | 10.143 |

## DISCUSSION

Paradigmatic responding is a specific language feature that apparently exists as a skill independent from reading or intellectual measures. The correlational relationships between paradigmatic language and reading or intellectual measures are nonsignificant and show a low percentage of common variance. The skills appear to be unique from one another. Commonality exists, but independence is assured in measurement. Though nonsignificant relationships exist among the variables, there is a tendency for paradigmatic language to form a more substantial relationship with intellectual rather than academic measures. Perhaps paradigmatic language does possibly reflect an overt facet of the cognitive process, not a facet of instruction. The levels of paradigmatic responding are observably existant and the relationship with the cognitive model seems measureable.

Training in paradigmatic language responding does not seem to enhance reading achievement in the areas of word meaning or paragraph reading when disabled readers are the subjects. No group mean differences were evident in the covariance analyses. Training at the sixth grade level may not be the necessary therapy to remediate reading problems. It may be possible that the reading and thinking habits of the age group are already structured and not readily amenable to change after training procedures of short duration. Perhaps paradigmatic training at a lower level, first or second grade, may be more productive. Training at the lower grade levels may provide the communication and thinking tools that provide learning of basic reading skills. Training at the sixth grade level, with disabled readers, may be too little too late.

Paradigmatic language responding is a skill that can be taught. The covariance analysis followed by Scheffe comparisons supports this conclusion. The experimental training group made significantly more paradigmatic responses than the other groups. Paradigmatic associations can be significantly taught. Growth is rapid and positive. However, transfer exercises to academic fields is possibly neded if this skill is to affect reading performance. Also the control, no training group, made more paradigmatic responses than the placebo group. A possible explanation for the retarding effect of the placebo treatment with important implications may be drawn. The nature of the placebo treatment, roughly analogous to classroom language enrichment exercises, may have confused the responding/thinking patterns of the subjects. Poem writing, story telling, listening to music for verbal stimulation, and word games may retard paradigmatic language processes. Nondifferentiated language training is possibly not beneficial to language growth. No training is apparently better in fostering language development than training with no observable goals. Instructional goals (classroom) should be established and projected prior to training.

Further investigations into the nature and effects of paradigmatic associations are needed. These investigations need to focus on training studies providing transfer to academic tasks; also the future training studies should possibly be conducted in younger than sixth grade groups.

## REFERENCES

Bickley, A. C., Dinnan, J. A., & Jones, J. P. Oral associates and reading readiness. In F. P. Greene (Ed.), *Reading: the right to participate*. National Reading Conference 20th Yearbook. Milwaukee: the Conference, 1971. Pp. 14-16.

Bickley, R. T., Bickley, A. C., & Cowart, H. Oral language responses and reading performances in the intermediate grades. In F. P. Greene (Ed.), *Reading: the right to participate.* National Reading Conference 20th Yearbook. Milwaukee: the Conefrence, 1971. Pp. 11-13.

Dinnan, J. A. Key to learning. *Journal of Reading Behavior,* 1971, *3,* 1-13.

Dinnan, J. A., Bickley, A. C., & Bickley, R. T. Reading achievement compared with oral language responses. *Journal of Reading,* in press.

Dinnan, J. A., Bickley, R. T., & Williams, J. C. Syntagmatic oral responses and SAT scores of college freshmen. In F. P. Greene (Ed.), *Reading: the right to participate.* National Reading Conference 20th Yearbook. Milwaukee: the Conference, 1971. Pp. 17-19.

Dinnan, J. A., Cowart, H., & Bickley, A. C. An analysis of semantic products of disadvantaged and advantaged second, third, and fourth grade good and poor readers. *Journal of Reading Behavior,* 1971, *3*(2), 22-26.

Entwisle, D., Forsyth, D., & Muuss, R. The syntactic-paradigmatic shift of childrens' word associations. *Journal of Verbal Learning and Verbal Behavior,* 1964, *3,* 19-29.

Ervin, S. Changes with age in verbal determinants of word association. *American Journal of Psychology,* 1961, *74,* 361-372.

Gunderson, D. *Language and reading.* Washington: Center for Applied Linguistics, 1970.

McNeil, D. A study of word association. *Journal of Verbal Learning and Verbal Behavior,* 1966, *5,* 548-557.

McNinch, G., Richmond, M., & Hornsby, J. Syntagamatic-paradigmatic shifting in word associations across grade levels. *Southern Journal of Educational Research,* 1970, *4,* 193-207.

Shen, S., & McNinch, G. The relationship between language associations and reading achievement in the eighth grade. *Southern Journal of Educational Research,* 1971, *5,* 40-50.

ALLAN ROY OHAVER
Harrisburg Area Community College

# A COMPARISON STUDY OF SEMANTIC AND SYNTACTIC CUEING BY LOW READING PERFORMANCE COLLEGE FRESHMEN

The general objective of this study was to investigate syntactic and semantic cueing, as used in oral reading by low reading performance college freshmen. Each subject had an overall test score at or below the thirty-fifth percentile on the national norms of the *Nelson-Denny Reading Test*. A survey of related psycholinguistic research revealed that syntax and semantics effected the perception of utterances (Fodor & Bever, 1965; Mehler & Carey, 1967; G. Miller, 1962; Miller & Isard, 1963; I. Miller, 1964; O'Neill, 1957). Other literature indicated that the same relationship should hold for a language user's comprehension of written expressions (Abrams, 1966; Goodman, 1965; Morton, 1964; Weber, 1968). A study was run (Ohaver, 1971) to examine these relationships.

## PROCEDURE

The subjects were 30 freshmen enrolled at Southern Colorado State College. One group of 15 was composed of subjects with an appreciably higher reading vocabulary subtext score on the *Nelson-Denny Reading Test*. An appreciable difference was defined as 12 or more percentile ranks between scores. The other group of 15 was composed of subjects who had an appreciably higher subtest score in reading comprehension than reading vocabulary.

Altogether one hundred and eighty-six freshmen met the following requirements for selection: They were "Predominant English Speakers," i.e. they could not carry on a meaningful conversation without recourse to English; they had a composite reading test score at or below the thirty-fifth percentile; and there was an appreciable difference betwen their vocabulary and reading comprehension subtest scores. Table 1 gives the subtest scores of the two groups on the *Nelson-Denny Reading Test* and the length of time it took the groups to read an expression-type instrument.

Table 1

A Descriptive Comparison of the Two Sample Groups

| Scores | Groups | Mean | Median | Range |
|---|---|---|---|---|
| Vocabulary[c] | High C[a] | 12.33 | 12 | 6–26 |
| | High V[b] | 37.26 | 37 | 14–83 |
| Comprehension[c] | High C | 34.60 | 35 | 21–57 |
| | High V | 10.40 | 9 | 2–30 |
| Reading Time[d] | High C | 8′40.9″ | 7′28″ | 5′39″–13′53″ |
| | High V | 8′37.5″ | 8′12″ | 4′53″–9′48″ |

[a] High C's are the Higher Comprehension Subjects.
[b] High V's are the Higher Vocabulary Subjects.
[c] Scores are percentile ranks.
[d] Times are minutes and seconds.

This instrument consisted of three sets of fifty typewritten expressions: Sentences, Semi-Grammatical Sentences (which were semantically anomalous), and Ungrammatical Strings. All one hundred and fifty expressions were randomly assembled and presented to the subjects on xerox copies. The same material had been previously presented to subjects on a tape recording by Miller and Isard (1963) to effectively demonstrate the effect of semantic and syntactic processing on the auditory perception of utterances. All multiple comparisons were by Friendman's Multi-Sample Test. All other comparisons were by Wilcoxon's *Rank Tests.*

The expression-type instrument originally grew out of the finding of Miller, Heise and Lichten (1959) that content words (nouns, verbs, adjectives and adverbs) were more intelligible when heard in the context of a grammatical sentence than when scrambled and heard in isolation. After reading Chomsky's *Syntactic Structures* (1957) and reviewing some related research, Miller (1962) concluded that an important aspect involved in the understanding of any sentence was the determination of the sentence's phrase structure. Within this theoretical framework Miller and Isard (1963) developed an instrument to aid in the determination of how a listener makes use of grammatical context as an aid in his identification of a content word.

The preset study, like the Miller and Isard investigation, was limited to syntactic and semantic cueing. Phonological cues were controlled by using the same words throughout. It was reasoned that when an "utterance" is constructed so that a listener can infer its phrase structure, he can use this structure to delimit his choice of content word alternatives, regardless of the content word's position in the utterance. Similarly it was deemed important to distinguish between a reader's use of syyntactic and semantic cueing. However, it must be noted that this study made no attempt to control for the pragmaic, or nonlinguistic information normally present in an ordinary spech transaction.

Originally the Semi-Grammatical Sentences were generated by taking five sentences which were semantically well-formed and had identical phrase structures, but were not necessarily syntactically equivalent. By substituting some of the words in one sentence for some in another Semi-Grammatical Sentences were produced as illustrated in the following example:

A witness signed the official legal document.
A jeweler appraised the glittering diamond earrings.
A magazine exposed the shocking political corruption.
A storm prevented the annual company picnic.
A knight slew the ferocious fire-breathing dragon.

Which when scrambled resulted in:

A witness appraised the shocking company dragon.
A jeweler exposed the annual fire-breathing document.
A magazine prevented the ferocious legal document.
A storm slew the official diamond corruption.
A knight signed the costly political picnic.

Although the sentences which resulted from this proces appear to be grammatically correct, the semantic cues normally assumed by a reader have been violated (Miller & Isard, 1963).

In a third category, Ungrammatical Strings, each set of five expressions was produced by a haphazard permutation of the position of the words to produce expressions as follows:

A legal glittering the exposed the picnic knight.
A diamond shocking the prevented dragon witness.
A political annual the slew document jeweler.
A company ferocious the signed earrings magazine.
A fire-breathing official the appraised corruption storm.

The one hundred and fifty expressions produced were then rearranged in a completely random manner. The first few lines of the expression-type instrument were as follows;

Tempers young rhythm ate secret the. Became lecture the bar deep wealth a. The cleverly disguised criminal fooled everybody. Bloom wildcats loudly yellow healthy. Fooled colored gently the restaurant cancer. A fire-breathing official the appraised corruption storm. Healthy young babies sleep soundly.

The subjects were scored both for the number of miscues made on the content words in each category, and for the number of corrections made on the content words in each category. Self-corrections of miscues were counted as corrections and not as miscues.

## RESULTS

The results of the miscues studied indicated that the Higher Comprehension subjects used both semantic and syntactic structure in oral reading, whereas the Higher Vocabulary subjects mainly used syntactic structure. This was concluded after extending Friedman's multi sample test in two directions according to procedures presented by Bradley (1948, p. 138-141) with each subject in effect providing the matched observations across the three expression types. The rather unusual results presented in Table 2 when further analyzed by Wilcoxon's Signed-Rank test (table 3) indicated that although the expression types used did have some effect on subjects' miscues it was the higher comprehension subjects who made significantly more use of the semantic component of the language.

Table 2

Results of Friedman's Multi-Sample Tests on Miscues

| Test Results | Rank Totals | | | | |
|---|---|---|---|---|---|
| | GS[a] | SGS[b] | USW[c] | S | Rows |
| Higher Vocabulary Subjects | 20.5 | 24.5 | 45.0 | 345.5• | 15 |
| Higher Comprehension Subjects | 20.5 | 24.5 | 45.0 | 345.5• | 15 |
| Main Effects of Expressions on Subjects | 20.5 | 24.5 | 45.0 | 345.5• | 15 |

[a] GS = Grammatical Sentences
[b] SGS = Semi-Grammatical Sentences
[c] USW = Ungrammatical Strings of Words
• $p < .001$

The results of the corrections studied indicated that the Higher Comprehension subjects made about the same number of corrections in each category; and that an increase in linguistic structure produced a corresponding decrease in corrections for the Higher Vocabulary Subjects. It should be noted that for these subjects who ranked in

Table 3

Results of Wilcoxon's Signed-Rank Tests on Miscues

| Test Results | W+ | n |
|---|---|---|
| Higher Vocabulary Subjects | 37.0 | 14 |
| Higher Comprehension Subjects | 24.5• | 14 |
| Main Effects of Expressions on Subjects | 23.0• | 14 |

• $p < .05$

approximately the lower third of the *Nelson-Denny* only the effects of syntax made a difference on corrections as indicated by the fact that although the main effects of expressions was significant across all three types (table 4) there was no significant difference between the main effects of Semi-Grammatical Sentences and Ungrammatical Strings (table 5). The Higher Comprehension group produced sig-

Table 4

Results of Friedman's Multi-Sample Tests on Corrections

| Test Results | Rank Totals | | | | |
|---|---|---|---|---|---|
| | GS[a] | SGS[b] | USW[c] | S | Rows |
| Higher Vocabulary Subjects | 22.5 | 28.5 | 39.0 | 139.0• | 15 |
| Higher Comprehension Subjects | 29.5 | 24.0 | 31.0 | 3.5 | 15 |
| Main Effects of Expressions on Subjects | 22.5 | 28.5 | 39.0 | 139.0• | 15 |

[a] GS = Grammatical Sentences
[b] SGS = Semi-Grammatical Sentences
[c] USW = Ungrammatical Strings of Words
• $p < .01$

Table 5

Results of Wilcoxon's Signed-Rank Tests on Corrections

| Test Results | W+ | n |
|---|---|---|
| Higher Vocabulary Subjects | 22.5• | 15 |
| Main Effects of Expressions on Subjects | 18.5 | 12 |

• $p < .025$

nificantly more miscues than the Higher Vocabulary group. There was no significant difference between the two groups in terms of total number of corrections produced.

Table 6
Wilcoxon Rank-Sum Test to Compare the Main Effects Between the Subject Groups

| Basis of Comparison | Wn | m=n |
|---|---|---|
| Miscues | 191.5* | 15 |
| Corrections | 201.5 | 15 |

* $p < .05$

From the results of this investigation, it appears that the subjects who are higher in reading comprehension are using a semantic element in reading that the other subjects are not using. What is not immediately evident is how both groups are using the syntactic element in reading. The first impression one might get is that they are both using it with equal skill, since they both read Semi-Grammatical Sentences better than Ungrammatical Strings of Words. However, if, as many authors indicate, reading comprehension is a language related process (Briggs, 1967; Lefevre, 1964; Rystrom, 1970) then the finding that it is the Higher Comprehension students who make significantly more miscues does not support this conclusion. Instead, the indication is that Higher Comprehension subjects are probably reading deep structure, and hence producing more miscues when they retransform (Goodman & Burke, 1969), to produce the meaning obtained. This finding would be in concord with Morton's earlier findings (1964).

Probably, the Higher Vocabulary subjects have become effective recoders, i.e., they are recoding not only from the grapheme to phoneme, but also have learned to recode the surface aspects of written syntactic structures without using the underlying structure. Properly used syntactic structures should lead to a semantic interpretation. According to Chomsky "the syntactic component of a grammar must specify, for each sentence, a *deep structure* that determines its semantic interpretation and a *surface structure* that determines its phonetic interpretation [1965, p. 16]."

The Higher Vocabulary subjects made the most corrections on the words in expressions that contained the least linguistic structure. This may indicate a greater concern with pronouncing, or at least knowing the correct word, than with grasping a sentence's meaning. These subjects apparently could not bring their vocabulary to bear on the reading task, because they had limited their choices to word meanings and syntactic structure. Clearly, if one is reading for meaning, it will be the total semantic meaning that determines his lexical item choices.

In contrast it appears that the primary limitation of the Higher Comprehension subjects was their restricted vocabulary. These subejcts needed more lexical items to process, since they appeared to be able to respond with material that satisfied both the syntactic and semantic requirements of an expression.

If, as this study indicates, the Higher Comprehension reader is operating as a language user in reading, then where has the reading process broken down for the Higher Vocabulary subjects? Is there a model which will suggest where the problem lies? Although these findings sem to be compatible with Goodman's (1970) model of reading, the model is quite complex, as is any model which takes into account all of the diverse aspects of reading. Fortunately, Wales and Marshall (1966) have constructed a simpler Schema of Linguistic Performance which will serve our purposes:

> first, it indicates the hypothesized order of processing linguistic information and suggests the most fruitful points of study; secondly, by comparison with alternative formulations empirical grounds for substantive claims can perhaps be tested; thirdly, it helps to keep as a central aim that linguistic performance is best considered as a system and avoids the proliferation of unrelated models which account merely for particular instances; . . . (p. 55).

Wales and Marshall, in constructing this schema, had in mind a continuous unidirectional process with sentences as the basic unit.

The apparent applicability of this schema will be clearer after a short discussion of each of its components. First, *preliminary recognition* is the incomplete sampling of the redundancy available in linguistic surface structure. Next, the *state of readiness* accounts for the effect of short term memory via a threshold mechanism, which is presumably regulated by situational expectancies and feedback instructions on the basis of sentential constraints and semantic coding. *Preliminary analysis of surface structure* is the stage where competent language users process their bits and pieces of surface structure to provide an input of the deep structure analyser. (However the Higher Vocabulary subjects studied in this investigation seem to have processed the entire surface structure, and proceeded from the third stage directly to the production of an oral output.) The *deep structure analyser* assigns relational notions upon a phrase-structure in such a way that its output is a limited set of items. It is at this stage that words are looked up in a semantic dictionary, which only operates on syntactically classified material. The *conceptual matrix* is not clearly differentiated from the semantic *interpretation* state. How-

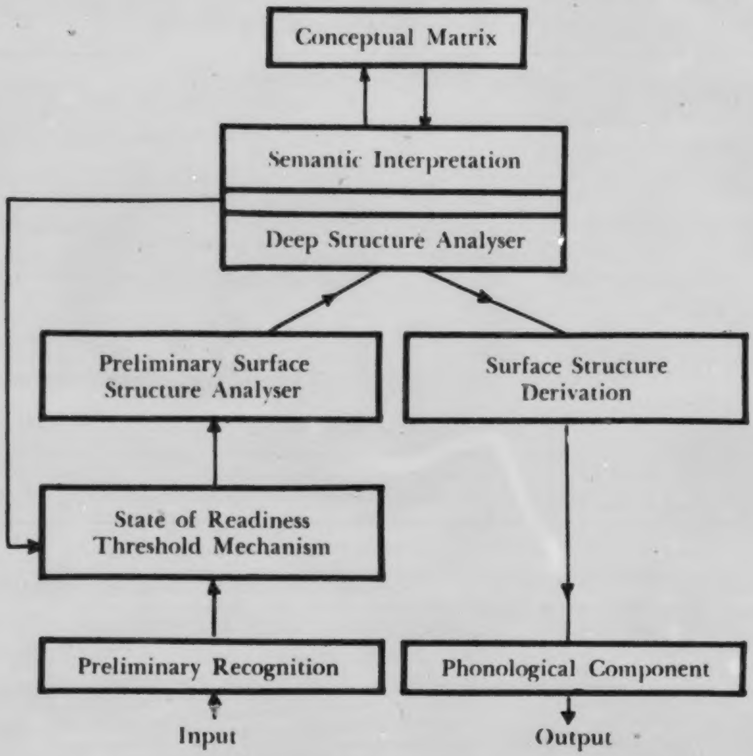

Conceptual Matrix
Semantic Interpretation
Deep Structure Analyser
Preliminary Surface Structure Analyser
Surface Structure Derivation
State of Readiness Threshold Mechanism
Preliminary Recognition
Phonological Component
Input
Output

**Fig. 1. Schema of Linguistic Performance. (Wales and Marshall, 1966)**

ever, it is only later in production, i.e., the last two stages, that all the *structural* information is processed.

Although one cannot clearly conclude from this study how the processing in one stage affects the other, one can hypothesize that an excessive amount of linguistic input to the *preliminary surface structure analyser* would create an overload that short circuits to the production stage. "Instead of going from print to meaning, they wind up with something that isn't meaning at all [Goodman, 1971, p. 462]."

## EPILOGUE

The objective findings of this study are interpreted as supporting the position that "Reading is not reading unless it involved some level of comprehension. [Smith, Goodman, & Meredith, 1970, p. 25]." During the last sixty years this position has in one form or another been involved in nearly every debate about reading. Unfortunately, much of the research related to reading has ignored Huey's warning: "With meanings, as with vocal utterance, the sentence-meaning is the natural unit, and smaller divisions considered apart from this are felt as *disjecta membra* [1908, p. 167]."

## REFERENCES

Abrams, A. G. The relation of listening and reading comprehension to skill in message structuralization. *Journal of Communication*, 1966, *16*, 116-125.

Bradley, J. V. *Distribution-free statistical tests*. Englewood Cliffs: Prentice-Hall, 1968.

Briggs, F. A. Implications of structures in language. In G. B. Schick & M. M. May (Eds.), *Junior college and adult reading programs—expanding fields*. National Reading Conference 16th Yearbook. Milwaukee: the Conference, 1967.

Chomsky, N. *Syntactic structure*. The Hague: Morton, 1957.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge: Massachusetts Institute of Technology Press, 1965.

Fodor, J. A., & Bever, T. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, *4*, 414-420.

Goodman, K. S. A linguistic study of cues and miscues in reading. *Elementary English Journal*, 1965, *42*, 639-644.

Goodman. K. S. Reading: a psycholinguistic guessing game. In H. Singer & Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, Del.: International Reading Association. 1970. Pp. 259-272.

Goodman. K. S. Decoding—from what to what? *Journal of Reading*, 1971, *14*, 455-462.

Goodman, K. S., & Burke, C. L. *A study of oral reading miscues that result in grammatical retransformations*. Project 7-E-219 final report of U.S. Office of Education. June 1969.

Huey, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908. (Republished: Cambridge: Massachusetts Institute of Technology Press, 1968.)

Lefevre, C. A. *Linguistics and the teaching of reading*. New York: McGraw-Hill, 1964.

Mehler, J., & Carey, P. Role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 335-338.

Miller, G. A. Some psychological studies of grammar. *American Psychologist*, 1962, *17*, 728-762.

Miller, G. A., Heise. G. A., & Lichten. W. The intelligibility of speech as a function of the context of the text materials. *Journal of Experimental Psychology*, 1951, *41*, 329-335.

Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, *2*, 217-228.

Miller, I. Perception of nonsense passages in relation to amount of information and speech-to-noise ratio. *Journal of Experimental Psychology*, 1957, *53*, 388-393.

Morton, J. A. A model for continuous language behavior. *Language and Speech*, 1964, *7*, 40-70.

Ohaver, A. R. A comparison study of semantic and syntactic cueing by low reading performance college freshmen. (Doctoral dissertation, New Mexico State University) Ann Arbor: University Microfilms, 1971. No. 71-28,050. Dissertation Abstracts 1971, *32*, 2668A.

O'Neill, J. J. Recognition of intelligibility test materials in context and isolation. *Journal of Speech and Hearing Disorders*, 1957, *22*, 87-90.

Rystrom. R. Toward defining comprehension: a first report. *Journal of Reading Behavior*. 1970, *2*, 56-74.

Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in the elementary school*. New York: Holt, Rinehart & Winton, 1970.

Wales, R. J., & Marshall, J. C. The organization of linguistic performance. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistic papers*. Edinburgh: University Press, 1966.

Weber, R. M. Study in oral reading errors. *Reading Research Quarterly*, 1968, *4*, 98-119.

ERNST VON GLASERSFELD
University of Georgia

# READING, UNDERSTANDING, AND CONCEPTUAL SITUATIONS

While most of you have been predominantly interested in the ways and means of teaching children to process and understand written language, I have spent my time trying to teach just that to a computer. But between children and computers there still is, fortunately for us all, a world of difference. While computers are very quick and very accurate, they manifest none of the self-reliance, independence, originality, or charm that so often encourage and stimulate us in our dealings with children. Still, there are similarities, especially if we confine ourselves to studying the acquisition of one specific skill.

There are overwhelming differences, too, between a seamstress and a sewing machine; yet, with regard to that one skill of sewing, viewed as a process that leads from a given input to a required output, there are similarities that are both interesting and instructive. Whoever it was who first hit upon the idea of mechanising the process of sewing was at once compelled to define, investigate, and analyse that process in a much more minute and accurate way than had ever been necessary before. Similarly, if we intend to make a machine understand sentences, we have to look much more closely at the process of understanding than seems necessary in the case of children. Hence it may not be quite unreasonable to expect that at least some of the insights that prove helpful in work with machines will be of use also in an attempt to facilitate or to speed up the learning process in children.

The use of the word "understanding" should make it clear that I am here not concerned with word recognition in speech or written material, nor with any of the sensory or perceptual activities involved. I am concerned exclusively with the processing of linguistic input that has already been recognized as such by the receiving organism.

Ziff (1970) says that we don't know what we are talking about when we use the verb "to understand". (Throughout this text double quotes are used to indicate that the marked item is to be taken as *word-token;* single quotes are used to indicate that the marked item is to be taken as *designatum,* i.e. as the concept or conceptual structure semantically linked to the word or word combination.) Ziff may be justified from a formal logical point of view. I would suggest, however, that, in practice, we have a fairly clear idea of what it is that a child has to learn in order to read and to comprehend or understan da sentence. The trouble is that linguists, almost without exception, have tended to study language as a ready-made observable phenomenon. As good scientists do in other branches, they have been intent upon discovering regularities, patterns, and rules that might help them to isolate, classify, and describe as much as possible of the phenomenon's immediately manifest character. Chomsky has made this point with admirable clarity:

> To avoid what has been a continuing misunderstanding, it is perhaps worth while to reiterate that a generative grammar is not a model for a speaker or hearer. . . . When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation (1965, p. 9).

However, there are still linguists and psycholinguists who would have us believe that a grammar based on the surface characteristics of words and word combinations, can supply us with a useful model for language acquisition, which is intended to include the process of language interpretation, or understanding, which the child has to acquire in order to read successfully. Fortunately, however, there are today also quite a few language analysts and psychologists—Piaget, Ceccato, Lenneberg, Hays, Schank, to mention only some—who have made it clear that, whatever a model of a language processing organism might eventually be like, it will have to be based on a plausible model of cognitive development.

## CONCEPTUAL STRUCTURES

This is not the place to describe fully a possible model of cognitive development. All I can hope to do within the restricted space at my disposal is to illustrate by means of a few examples how given pieces of language may be linked to certain cognitive or conceptual structures and that such links are crucial in the process we call understanding.

The early stages of a child's cognitive development were summarily hypothesised by Piaget (1926) almost half a century ago. There seems to be little doubt today that the construction of relatively durable object concepts *(object conservation)*, and of some basic relational concepts with which to connect the static objects, are among the very first cognitive achievements of the child. It also seems very plausible now that the relational concepts are first derived from the various efforts the child makes to coordinate his own motor sequences to the visual and tactual data patterns it has come to hold constant in the form of object representations (Piaget's *sensory-motor* stage).

Although an infant of a few months of age is, of course, still very far from having a conscious concept of 'activity' and of 'causation', the motor patterns it learns to implement (e.g. to reach for, grasp, and shake a rattle) already contain elements which, at a much later stage, will be categorized as 'causative agent', 'actor', 'activity', 'object', and 'caused effect' or 'result'. And these are the very elements of which the kind of situation is constituted that the child learns to express in language by means of a subject, a verb, and various complements. In order to achieve this linguistic expression, the child must learn to master the semantic connections by means of which his language designates individual conceptual items (i.e. 'objects' which speakers normally hold constant) as well as relational concepts (i.e. the relations which link individual conceptual items in a given situation). As we shall see, this dichotomy between individual static items and the relational concepts that link them is only very approximately reflected on the linguistic level by the grammarian's division between words (lexical items) and sentence structure (syntax).

The child will have to learn, for instance, that sentences 1 through 7 must all be referred to one and the same conceptual situation:

1. The baby shakes his rattle.
2. The rattle is shaken by the baby.
3. The rattle shakes and rattles.
4. The rattle rattles.
5. The baby rattles with his rattle.
6. The baby rattles.
7. The baby's rattling (irritates father).

It would be of little help to the child (who, we can safely assume, is not much concerned with grammatical description) if a linguist were to tell it that sentence 2 can be generated by means of transformation and substitution rules from sentence 1, for the linguist would be quite unable to provide any further linguistic rules for the generation of sentences 4, 5, 6, and 7. It is, indeed, difficult to see how anyone could understand (or, for that matter, produce) such a set of sentences unless he had developed the conceptual network that enables him to assign, to the items designated by the individual words, the roles they are playing in the event to which the given sentence refers. Moreover, some of the sentences do not even contain words for all the items that, in fact, play important roles in the event. In sentences 3 and 4 the causative agent is in no way designated. In sentences 6 and 7 it is the item that actually performs the activity of "rattling" (i.e. "the rattle") that remains wholly implicit. The examples also demonstrate that a grammatical parsing that yields parts of speech does not help us with the assignation of roles.

Thus, the conclusion seems inescapable that in order to understand even relatively simple sentences we must integrate such information as the sentences yield, with information that we have to provide ourselves. This additional information need not be linguistic information, in the sense that it was obtained from or through the use of the language. It frequently is, and sometimes can only be, experiential information.

If we look at the above sentences, we can see very quickly that the experiential information required to understand them, consists primarily of the following static and relational concepts:

A. a causative agent (e.g. designated by "baby") which causes
B. a non-causative item (e.g. designated by "a rattle") to perform
C. an activity (e.g. designated by "to shake") which, in turn, causes
D. an activity (e.g. designated by "to rattle").

Apart from this information at least four pieces of lexical information are also relevant:

E. that "baby" designates the type of item that can function as 'causative agent';
F. that "rattle" designates the type of item that can enact both the activities designated by "to shake" and "to rattle";
G. that "to shake" designates an activity that involves repetitive 'change of place';
H. and that "to rattle" designates an activity that involves the emission of a specific repetitive type of noise.

In time, as experiential knowledge increases, the conceptual structure will become much more detailed: the causal nexus between A and C will be specified as mechanical causation (motion imparted by mechanical force, e.g. through prehensile contact); the production of noise will be specified as caused by the intermittent contact of two items; the particular acting object "rattle" B will be specified as a purposive instrument consistin of at least two items that can make and break contact with one another; and so on.

Only when these conceptual refinements (and others pertaining to other situations) have been incorporated, will the child be in a position to understand sentences of the kind:

8. The shutters were rattling in the wind.
9. The old pump was rattling in the yard.
10. The prisoner was rattling his chains.
11. The cart rattled down the hill.

It is easy to see that the determination of grammatical subject and object is of little help in the process of understanding, if by "understanding" we mean the correct assignation of roles in the described event. In fact, none of the things we can possibly find to say about the syntactic structure of these sentences tells us with any certainty which linguistic item designates the causative agent and which the performing actor in the conceptual situation designated by the verb "to rattle". What the syntactic structure does tell us in a good many cases, is which of the linguistic constituents of the sentence designates the particular conceptual situation to which the interpreter (or reader) must turn for such further information as he requires to achieve his interpretation. Both "rattle" and "rattles", for instance, may designate either the activity or the acting item. Their place in the word sequence helps us to decide whether, in the sntence at hand, they designate the dynamic conceptual situation or the static concept that happens to function as an acting item.

In short, from the words and their sequential arrangement in the English phrase or sentence we glean a certain number of semantic and syntactic indications, but in very many cases these indications are insufficient to reconstruct the actual situation to which the piece of language refers. A full reconstruction is then possible only if we map the gleaned indications onto a pre-existing network of conceptual items and relations.

Let us look at other examples:

12. Charley cuts bread very well
13. This knife cuts bread very well
14. This bread cuts very well

We could scarcely understand these sentences if we could not resort to a conceptual situation that is the same for all three because it is determined by the activity designated by "to cut" and not by what happens to be grammatical subject or object on the linguistic level.

Only when we are able to reconcile what we already know about the single items designed by the words of the sentence with the specific requirements of a known conceptual situation, can we say that we have understood the sentence. Each item, as it were, has to be fitted into a specific slot of the conceptual network. Fortunately, it is relatively rare that a sentence specifies several items that could be fitted equally well into one and the same slot (or into the slots of several different conceptual situations).

The assignation of roles is by no means the only interpretive problem that has to be decided in this way. A pre-established network is no less indispensable in determining to which item a modifier is to be related. In examples 15 through 22 the formal structure of the sentence does not tell us to which item the property designated by the last word of the string is to be applied; nor does it always tell us whether this last word, in the case at hand, designates a property or something else.

15. He climbed the hill fast.
16. He bent the wire down.
17. He bent the wire cold.
18. He sliced the ham thin.
19. He bought his car new.
20. He painted the car silver.
21. He offered the client silver.
22. He finished the race exhausted.

With the exception of "new", all the terminal words in these sentences, taken by themselves as lexical items, have more than one potential function in ordinary usage. "Thin" and "exhausted" can function as verbs. "Cold" can function as noun. "Fast" and "silver" can function both as noun and as verb. "Down" can function as verb, as noun, and as preposition. Besides, "fast" and "silver" are ambiguous also in one of their grammatical functions. The first, in its function as modifier, can be roughly synonymous to "speedy", to "tight", or to "indelible"; the second, in its function as noun can designate either a material or a color. (One could go on adding to this list of ambiguities, but the list is quite long enough to illustrate the points I want to make.)

Given this proliferation of potential functions, how does it come about that the reader, in practice, has very little, if any, doubt as to what the terminal word means and to what other item its modifying function is to be applied? Let me try to outline a possible procedure.

In sentence 15 "climbed", following directly upon "he", can be identified with near certainty as the personal form of a verb. As such it provides the central item of a dynamic conceptual situation. In fact, it designates an activity that involves a 'change of place' in an upward direction (unless there is the specification "down", which would invert the direction). Since in our example there follows the word "hill", which designates an item that fits very well into the slot provided by the second place, or place of arrival, in the 'change of place' structure, we are inclined to expect that the designation of the last word of the sentence will fit into the conceptual structure we have already built up, rather than that it will force us to reorganize the preceding items into a different conceptual structure. This expectation is, of course, based on our previous experience with the English language and has been fairly consistently reinforced by the fact that, in English, the word or word combination that designates the central conceptual structure tends to come rather early in the sentence. (For a detailed analysis of the different levels of expectation involved in the process of understanding see Schank, 1971.) Then, when we actually come upon the last word of the sentence, "fast" in our example, we do not even consider the potential activity-designation of that word because neither would the central activity around which we have already built a conceptual situation allow for a secondary activity to be added at this point, nor would we have a suitable 'acting object' that could be conceived as enacting this secondary activity. (This would be quite different if the sentence were something like: "He saw the monks fast" or "He knew the monks fast," where the conceptual structures designated by "to see" and "to know" do provide a slot for a second activity.) As to the potential noun-designation of "fast", this could become operative only if it provided us with a conceptual item that we could readily join to the item designated by "hill" without the latter losing its fitness for the slot we have called 'place of arrival.' (This would be the case if the last word were "side", making "hill side", or "track", making "hill track", but with the noun "fast" it does not seem plausible.) Thus we are left with the modifier-designations of "fast", and since the one that has to do with speed is overwhelmingly probable in conceptual structure that clusters around a 'change of place' or 'motion', we apply the modification to the activity or, as a rammarian would say, we apply it adverbially.

I should not like to strain your patience by going through an hypothetical interpretive procedure for the remaining seven examples. Let me merely point out the variety of ways in which the designatum of the last word has to be linked to other items in the diverse situations.

In example 16 "down" specifies the direction of the acting object at the terminal point of the activity.

In example 17 "cold" specifies a condition of the acting object before and during the activity.

In example 18 "thin" specifies a property of the results of the specific transformative activity, i.e. a property of the *slices* into which the ham is transformed.

In example 19 "new" specifies a condition of one of four items involved in the activity of 'reciprocal transfer' designated by "to buy" (the buyer, the seller, the item acquired by the buyer, and the item, e.g. money, received, or to be received, by the seller).

In example 20 "silver" specifies a property of the transformed item, i.e. a property that is the result of the specific transformative activity.

In example 21 "silver" designates the acting object of an activity of 'transfer' that is merely projected by the activity designated by "to offer".

In example 22 "exhausted" specifies a condition of the acting object at the terminal point of a 'change of state', i.e. the change from 'activity' to 'non-activity' designated by "to finish".

## A CONCEPTUAL UNIVERSE

These rough sketches of conceptual situations by no means exhaust the relational network that could be uncovered in each one of them. But, rough as they are, they serve well enough to illustrate the amount and the complexity of well-specified notions (as to the semantic structure of lexical items and the relations into which, potentially, they can be fitted)—notions which are indispensable to the reader if he is to understand linguistic expressions, even if neither syntax nor lexical items, in the expressions are in any way out of the ordinary.

The number and variety of these notions and, beyond that, the question as to how they may be accumulated and integrated into a relatively coherent conceptual universe, should now help us to bring into focus one of the big and as yet unbridgeable differences between the attempts to teach reading to a child and to a computer. The child, when it reaches the age at which reading is usually contemplated, has had four or five years of living experience during which he has built up at least a skeleton of his conceptual universe and he has had at least two or three years of linguistic experience supplying a multitude of information with which to extend and to substantiate that skeletal structure. A computer has had none of that. Unless we can supply experience ready-made, or can program an efficient self-learning procedure, the day when a computer will read and understand ordinary English is not likely to come.

Though the type of analysis of conceptual situations I have tried to outline in these pages has sprung from work with computers, it does, I believe, throw a little light on some aspects of language and language skills that many linguists leave in the dark. As long as the analysis of a piece of language yields merely rules for the transformation of one formulation into another and rules for the generation of sentences from a deep structure whose components are no less linguistic than the generated result, the corpus of these rules will be of very little help to either the student or the teacher of a linguistic skill such as reading. For the crucial step in the interpretation of written (or, indeed, of spoken) language is not the step from one linguistic structure to another, but the step from a linguistic expression to a non-linguistic conceptual representation. As I have tried to show, the conceptual situation is often only summarily specified by the explicit linguistic input. Understanding, therefore, requires a varying amount of additional information that necessarily has to be gathered from a pre-established conceptual network.

The failure to understand a given sentence may, thus, not always be due to a lack of linguistic competence. It could be caused equally well by a gap in the reader's conceptual universe. Although this has, of course, been well known insofar as the reader's vocabulary is concerned, we have only recently become fully aware of the fact that the correct relational interpretation of sentence structure frequently depends on relational information which the syntax of the sentence itself does not supply.

No less important than this additional relational information is the creation, in the reader, of certain expectations concerning the content of parts of the sentence (or, indeed, text) that have not yet been input or processed. Again there can be little doubt that these expectations are generated when the reader maps information gathered from single words and phrases onto much larger pre-existing conceptual structures.

There are, then, at least two ways—related, but rather different in function—in which the reader's pre-established conceptual network is crucial to his understanding of a piece of language. If this is so, it may be worth while to examine existing methods of reading instruction and, especially, the remedial strategies designed for backward readers, in order to see whether these methods and strategies could not be improved from this specific point of view, i.e. whether they are, in fact, as efficacious as they could be with a view to developing the student's conceptual universe.

## REFERENCES

Ceccato, S. Prima lettera ad un amico. Methodos, 1956, *8* (29/30), 139-156.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Mass.: M.I.T. Press, 1965.

Hays, D. G. Linguistic foundations of a theory of content analysis. In G. Gerbner, O. Holsti, K. Krippendorff, W. Paisley, and P. Stone (Eds.), *The analysis of communication content*. New York: Wiley, 1969.

Lenneberg, E. H. Of language knowledge, apes, and brains. *Journal of Psycholinguistic Research*, 1971, *1*, 1-19.

Piaget, J. *The language and thought of the child*. New York: Harcourt Brace, 1926.

Schank, R. C. Intention, memory, and computer understanding. *Research Report* STAN-CS-71-195. Stanford, Calif.: Computer Science Department, Stanford University, 1971.

von Glasersfeld, E. Semantics and the syntactic classification of words. Paper presented at the International Conference on Computational Linguistics, Sanga Saby, Sweden: 1969.

von Glasersfeld, E. Semantic analysis in terms of conceptual situations. *Linguistics*, 1971, in press.

Ziff, P. Understanding. In J. L. Cowan (Ed.), *Studies in thought and language*. Tucson: The University of Arizona Press, 1970.

NICHOLAS J. ANASTASIOW
Indiana University

# MEASURING CHANGE–A TIME DIMENSIONAL PROBLEM

This paper concerns some very simple issues. In so-saying the author recognizes that rather than being clarifying he may end up being simplistic. The issue to which this paper is addressed is the measurement of a child's or a group's current status in relation to a set of skills or abilities. In the case to be considered at this conference–reading. It will be proposed that the problem of measuring growth involves not only the traditional issues of reliability and validity i.e., selection, placement (predictions and decisions), content and the most elusive and ignored of all construct validity, but measurement also involves what and whom will be measured, at what critical points in time will you measure, and the use to which the measurement is aimed.

Measurement using normed reference scores often imply standards to teachers and parents. Scores that imply standards can be misleading as a definition of the word "standard" used as an adjective means:

(a) conforming to established custom, measure, or style
(b) serving as a test or measure for others

Used as a noun standard means:

(a) an established measure or model
(b) an upright timber used as a support in a building
(c) a flag or emblem

The thesis of this paper can be reduced to two major issues; one, there is an enormous national pressure to design and use measurement instruments to assess change; secondly, those who use measurement instruments to assess change are in danger of being scientifically, psychologically, and ethically involved in very questionable and unsupportable practices.

The misuse of tests or measuring instruments often occurs when tests are used to evaluate a program rather than assessing an individual or a group's current status. Using tests in this way is based on an erroneous assumption that gains can be measured utilizing our current techniques. Even the most impressive of statistical models, to my mind, does not handle the problems inherent in gain scores.

These issues have been ably discussed elsewhere by Stake & Wardrop (1971), Lennon (1971), in reference to performance contracting and by Thorndike (1971), in reference to regression effects. Regression effects can be handled statistically and if the score differences are large enough, they can be ignored altogether.

Psychologically, however, the use of gain scores to make inference as to causal relationship between current status and the treatment in the classroom is probably the most tenuous of assumptions as follows:

> General propositions embodying descriptive concepts must in the end be verified by systematic observation and the procedures used to gather those observations must be given a full observational description in order to make the report useful. One shows the construct is constant with the evidence; one does not show that the construct represents something that exists correctly (Cronbach, 1971, 482).

However, as one of the speakers at this conference has said there is something so appealing about gain scores. However, gain in weight, for example, may be caused by overindulgence, malfunctioning thyroid or cancer to name but a few. Gains in cognitive functioning may be through punitive, demeaning teachers as Spaulding (1963) demonstrated or through drill and practice procedures with its all too well known, but ignored, forgetting curve.

What is being said is that although measurement and evaluation are separate activities, evaluation cannot occur without measurement and it is *measurement* use in evaluation that, in my mind, is one of its major sources of difficulty.

Let us consider an hypothetical example based on one of the author's studies reported in full elsewhere by Fischler and Anastasiow (1965). A school district engaged in training teachers, employed in in-service procedures to implement inquiry techniques in elementary classrooms. Utilizing Flander's scale, teachers were observed before and after treatment and change in teachers' question asking and teacher lecture was demonstrated to be statistically significant following treatment, fortunately in the desired direction. That is, teachers engaged in more question asking and less lecture. The problem is attributing change of complex behaviors to the treatment. As Cronbach (1971) has stated, "it is important for testers to recognize that they use inferences and for them to scrutinize dispassionately the underlying evidence and reasoning [p. 483]."

If a teacher asks more questions and talks less, did she gain a skill from the treatment procedure, learn how to use a skill already possessed, or defer to the authority figure of the study and adopt his procedures to meet criterion performance. The assumption of causal relationships between treatment and performance is a weak nomological net—and in all probability is as scientifically unsafe as confusing learning and performance.

The answer to the above dilema may be that the evaluator really doesn't care whether the nomological net is tight—if he has been able to demonstrate that teachers or children will perform the desired behaviors and he is willing to restrict his generalization of the results. Similarly measurement *should not concern itself with change but the performance of the desired behaviors that define criterion performance from which traits can be inferred.*

As Cronbach & Furby (1970) summarize their position—"It appears that investigators who ask questions regarding gain scores would ordinarily be better advised to frame their questions in other ways [p. 80]." Or as Lord (1967) has stated, "There is simply no logical or statistical procedure that can be counted on to make proper allowances for uncontrolled pre-existing differences between groups [p. 305]."

Thus, treatment effects, in a multi-dimensional learning situation cannot be compared. What solution then can one have for those who wish to evaluate programs' effectiveness?

Anastasiow (1969) has shown that matching groups by similar stratified social class membership does much to control for extraneous group effects. Anastasiow (1968) has also shown that in an experimental reading program children in the new program excelled children on the criterion measures that were designed to measure the goals of the new program, whereas the comparison group excelled the experimental on measures designed to meet the goals of the control group. That is, "There is no need to measure change if testing the null hypotheses that two treatments have same effect—the essential question is do the post-test scores vary from group to group and the measure is a suitable dependent measure [Cronbach and Furby, 1970, p. 78]." In the study stated above, it is clear that the dependent measures selected, greatly modified the results.

In addition the assertion is made that the criterion reference points to be assessed change during the course of the program's evalution. Specifically the time dimension of measuring would follow somewhat the following 14 questions.

Only at point 11 would formal assessment be conducted and this measurement would involve teacher behavior. The data would be used to revise, continue, or recycle back to any of the first 10 points through

## FOURTEEN QUESTIONS TO ASK TO IMPROVE YOUR PROGRAM

| | |
|---|---|
| 1. What do you want to do? | Specific Goals |
| 2. Why do you want to do it? | General Goals |
| 3. How will you do it? | Philosophy<br>Curriculum Orientation<br>Psychology of Learning |
| 4. To whom will you do it? | Description of the Child Population |
| 5. In what way does the children you've selected modify what you want to | Nature of the handicap—Knowledge of Child Development (i.e., age of child) |
| 6. Who will help you do it? do? | Supporting Services (school psychologists, speech therapists, etc.) |
| 7. Who will do it? | Teachers, aides, parents as teachers |
| 8. What will they (teachers, etc.) do? | Content (Behavioral Objectives) |
| 9. How will they (teachers, etc.) do it? | Strategies for carrying out program teaching, community contacts, parent education |
| 10. How will they (teachers, etc.) learn how to do it? | In-Service Training or Pre-Service |
| 11. How will you know if they (teachers, etc.) are doing it? | Observation of teachers, aides, or parents. Questionnaires |
| 12. If they (teachers) aren't doing it, can you figure out why? | – Inappropriate Strategies<br>– Lack of In-service Training (inadequate or inappropriate)<br>– Poorly stated goals<br>– Poorly constructed curriculum<br>– Program doesn't fit objectives<br>– Lack of space<br>– Kids are different<br>– Lack of consultant help<br>– Insufficient resources |
| 13. When teachers have implemented the program, how do you know if children are making gains? | Pupil Growth<br>– Anecdotal records<br>– Pre-post measures<br>– Child as control<br>– Base line data<br>– Observational instruments<br>– Interviews<br>– Achievement and intellectual tests |
| 14. Finally, if it (your overall program) does work—how will you tell others? | Demonstration<br>Dissemination |

the solution of question 12. Only at question 13 are typical measures of reading comprehension appropriate. Criterion reference measures which are referenced to a curriculum sequence are more likely to serve the decision making process than other norm derived scores.

Let us close with some ridiculous but timely examples. Child X has otbained a perfect score on the STEP Reading Test 4-6 in 4th grade at the beginning of an evaluation of a modified Joplin Program. Does she have nothing more to learn? Child B, a first grader, is at the 99th percentile on the test by responding correctly to 39 out of 75 items. Does she need further instruction?

The dilema cannot be simply resolved by throwing away change score but *avoiding to attribute to score differences surplus and excess meaning*. In planning instruction one desires entry and exit behaviors but our current measures contain no true zero point and change at one level may contain quite different meaning than at another.

The current answer provided by Glaser and Nitko (1971) is to provide detailed diagnosis of students learning habits, attitude, achievement, skills and cognitive style and as the basis of this analyses of the student's characteristics guide him through course tailored for him.

An anecdote may help make the point. A child was asked by a teacher to say his times tables. He said, "la da da da da, la da di da di." The teacher interrupted and asked what he was doing. He said "I know the tune but I don't know the words [Cole, Gay, Glick, and Sharp, 1971, p. 17]."

As with measuring change, we have the problem but have not developed the appropriate methodology to do so.

## REFERENCES

Anastasiow, N. J. Progress report: evaluation of the story telling program. Palo Alto (Calif.) Unified School District Research Department, 1968.

Anastasiow, N. J. Social class as a matching variable in school district research. *Psychology in the School*, Oct. 1969, *6*, 396.

Cole, M., Gay, J., Glick, J., & Sharp, D. *The cultural context of learning and thinking: an exploration in experimental anthropology*. New York: Basic Books, 1971.

Cronbach, L. J. Test validation. In R. L. Thorndike (Ed.), *Educational Measurement*. (2nd. ed.) Washington, D.C.: American Council on Education, 1971. Pp. 443-03.

Cronbach, J. J., & Furby, L. How should we measure "change"—or should we? *Psychological Bulletin*, July 1970, *74*, 68-80.

Fischler, A., & Anastasiow, N. J. In-service education in science (a pilot)—the "school within a school." *Journal of Research in Science Teaching*, Dec. 1965, *3*, 280-285.

Glasser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), *Educational Measurement*. (2nd. ed.) Washington, D. C.: American Council of Education, 1971. Pp. 625-670.

Lennon, R. Y. Accountability and performance contracting. Paper presented at the meeting of the American Educational Research Association, New York, Feb. 1971.

Lord, F. M. A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 1967, *68*, 304-305.

Spaulding, R. L. Achievement, creativity and self-concept correlates of teacher-upil transactions in elementary schools. U. S. Department of Education Cooperative Research Report No. 1352. Urbana: University of Illinois, 1963.

Stake, R. E. Testing hazards in performance contracting. *Phi Delta Kappan*, 1971, *52*, 583-589.

Stake, R. E., & Wardrop, J. L. Gain score errors in performance contracting. *Reading Research Quarterly*, 1971, *6*, 323-325. Also in *Research in the Teaching of English*, 1971, *5*, 226-229.

Thorndike, R. L. Reproducing the test. In R. L. Thorndike (Ed.), *Educational Measurement*. (2nd. ed.) Washington, D. C., American Council on Education, 1971. Pp. 160-187.

**FREDERICK B. DAVIS**
University of Pennsylvania

# A COMPARISON OF STATISTICAL MODELS USED IN THE MEASUREMENT OF CHANGE

Estimation of change can be accomplished in a number of ways but the statistical models to be used fall mainly into four categories: raw-score estimates of change; regressed-score estimates of change; estimates of true change; estimates of residual change.

The first two equations in the left-hand column of Table 1 yield crude raw-score estimates of change for individuals. The standard errors of measurement for these estimates are shown in the right-hand column. The most serious fault of estimates of change expressed as differences between raw scores or between averages of two or more raw scores is that they systematically *overestimate* change for individuals who obtain pretest scores below the mean of the age or grade group in which the individual properly belongs and systematically *underestimate* change for individuals who obtain pretest scores above the mean of that group. Furthermore, the under-or overestimation varies differentially, becoming larger as the individual's score deviates farther from the mean. It is easy to see, therefore, that if pupils with the lowest scores on a pretest in reading are selected for remedial work, spurious overestimates of gains will be obtained if equations 1 or 2 are used. Because the average scores used in equation 2 are more reliable than single scores on the same test, the spurious overestimation of change is less marked than if equation 1 is used. Since it does not become negligible until the reliability coefficient of the pretest is close to unity, the use of raw-score individual measures of change is not recommended.

## Table 1
## ESTIMATION OF DIFFERENCES BETWEEN EQUIVALENT SCORES ON FORMS OF THE SAME TEST ADMINISTERED TO THE SAME INDIVIDUAL OR GROUP

| Estimation of Change | |
|---|---|
| *Raw-Score Point Estimate* | *Standard Error of Measurement of Point Estimate* |
| 1. $C = B - A$ | $s_{meas\ C} = \sqrt{s^2_{meas\ B} + s^2_{meas\ A}}$ |
| 2. $C'' = \frac{B + D}{2} - \frac{E + F}{2}$ | $s_{meas\ C''} = \frac{1}{2}\sqrt{s^2_{meas\ B} + s^2_{meas\ D} + s^2_{meas\ E} + s^2_{meas\ F}}$ |
| 3. $\bar{C}s = \bar{B}s - \bar{A}s$ | $s_{meas\ \bar{C}s} = \frac{s_{meas\ C}}{\sqrt{Ns}}$ |
| 4. $\bar{C}s'' = \frac{\bar{B}s + \bar{D}s}{2} - \frac{\bar{E}s + \bar{F}s}{2}$ | $s_{meas\ \bar{C}s''} = \frac{s_{meas\ \bar{C}''}}{\sqrt{Ns}}$ |
| *Regressed-Score Point Estimate* | |
| 5. $C' = B - E$ | $s_{meas\ C'} = \sqrt{s^2_{meas\ B} + s^2_{meas\ E}}$ |
| 6. $\bar{C}s' = \bar{B}s - \bar{E}s$ | $s_{meas\ \bar{C}s'} = \frac{s_{meas\ C'}}{\sqrt{Ns}}$ |
| 7. $Cr = r_{BB'}B - r_{AA'}A + L$ | $s_{meas\ Cr} = \sqrt{r^2_{BB'}s^2_{meas\ B} + r_{AA'}s^2_{meas\ A}}$ |
| 8. $\bar{C}rs = r_{BB'}\bar{B}s - r_{AA'}\bar{A}s + L$ | $s_{meas\ \bar{C}rs} = \frac{s_{meas\ Cr}}{\sqrt{Ns}}$ |
| *Point Estimate of True Change* | |
| 9. $\hat{C} = W_B B + W_A A + K$ | $s_{meas\ \hat{C}} = \sqrt{W^2_B s^2_{meas\ B} + W^2_A s^2_{meas\ A}}$ |
| 10. $\hat{\bar{C}}s = W_B \bar{B}s + W_A \bar{A}s + K$ | $s_{meas\ \hat{\bar{C}}s} = \frac{s_{meas\ C}}{\sqrt{Ns}}$ |
| **Estimation of Residual Change** | |
| *Point Estimation of Deviation From Predicted Change* | *Standard Error of Measurement of Point Estimate* |
| 11. $\hat{R} = B - \bar{B} - W''(A - \bar{A})$ | $s_{meas\ \hat{R}} = \sqrt{s^2_{meas\ B} + W''^2 s^2_{meas\ A}}$ |
| 12. $\hat{R}' = B - B - W'(A - A)$ | $s_{meas\ \hat{R}} = \sqrt{s^2_{meas\ B} + W'^2 s^2_{meas\ A}}$ |

where: $A$ = individual score on Form A (pretest)
$B$ = individual score on Form B posttest)
$C$ = crude point estimate of individual change
$C'$ = crude point estimate of individual change
$C''$ = crude point estimate of individual change
$D$ = individual score on Form D (posttest)
$E$ = individual score on Form E (pretest)
$F$ = individual score on Form F (pretest)
$\bar{A}$ = mean score on Form A in a defined age or grade group
$\bar{A}s$ = mean score on Form A in a nonrandom subsample of the group

$\bar{B}$ = mean score on Form B in a defined age or grade group
$\overline{Bs}$ = mean score on Form B in a nonrandom subsample of the group
$\overline{Es}$ = mean score on Form E in a nonrandom subsample of the group
$\overline{Cs}$ = crude point estimate of mean change in a nonrandom subsample of the group
$\overline{Cs}'$ = crude point estimate of mean change in a nonrandom subsample of the group
$\overline{Cs}''$ = crude point estimate of mean change in nonrandom subsample of the group
$\hat{C}$ = point estimate of individual true change
$\hat{\overline{Cs}}$ = point estimate of mean true change in a nonrandom subsample of group
R = point estimate of residual change (deviation from predicted change)
$\hat{R}'$ = point estimate of residual change (deviation from predicted true change)

$$K = \bar{B} - \bar{A} - W_B \bar{B} - W_A \bar{A}$$

$$L = \bar{B} - \bar{A} - r_{BB'} \bar{B} + r_{AA'} \bar{A}$$

$$W_A = \frac{r_{AB} s_B (1 - r_{BB'}) - s_A (r_{AA'} - r^2_{AB}}{s_A (1 - r^2_{AB})}$$

$$W_B = \frac{s_B (r_{BB'} - r^2_{AB}) - r_{AB} s_A (1 - r_{AA'})}{s_B (1 - r^2_{AB})}$$

$$W' = \frac{r_{AB} s_B}{r_{AA'} s_A}$$

$$W'' = \frac{r_{AB} s_B}{s_A}$$

Equations 3 and 4 apply if average scores of nonrandom subsamples of an age or grade group that has been pretested are used to estimate change. For example, a remedial-reading class might be formed for the 15 lowest-scoring pupils out of 200 pretested in reading at the beginning of grade 7. If the average pretest score of this group is subtracted from their average posttest score on an equivalent form of the test, the mean change of the group will be overestimated. Consquently, use of equations 3 and 4 for estimating change in a nonrandom subsample of a group is not recommended.

At this point it may be appropriate to explain why raw-score estimates of change are likely to be either under- or overestimates of true changes. When a pretest is given, the resulting scores are subject to random errors of measurement whenever the test has less than perfect reliability. In practice, this is always the case. Since the errors of measurement are random, they would, in the long run, cancel themselves out over all examinees and result in a mean error score of zero. However, the errors of measurement in the raw scores of the lowest-scoring pupils in a group are more likely to be negative errors of measurement than positive errors of measurement. Conversely, the errors of measurement in the raw scores of the highest-scoring pupils in a group are more likely to be positive errors of measurement than negative errors of measurement. The results derive from the fact that examinees whose scores happen by chance to include negative errors obtain lower scores than examinees of the same true ability whose scores happen to include positive errors; similarly, examinees whose scores happen to include positive errors obtain higher scores than those of the same true ability whose scores happen to include negative errors.

Now suppose that we choose out of a group of 200 pupils, the 20 pupils who obtained the lowest pretest scores. Their raw scores will, as noted above, include more negative errors than positive errors. If we give the man equivalent form as a posttest (regardless of whether any treatment is provided between the pretest and posttest), the errors of measurement in their raw scores can be expected to split equally among negative and positive errors. This decrease in the proportion of negative errors in the scores from pretest to posttest tends to cause the individual scores (and their average, of course) to rise toward the mean of the entire group of 200 pupils. In other words, nonrandom selection (of the 20 lowest-scoring pupils) leads to regression to the mean. Unless this regression is taken into account in the measurement of change, a gain or loss that is caused merely by chance may be interpreted as a true gain or loss. Fortunately, we have a variety of ways of taking it into account, the simplest of which is discussed next.

Use of a regressed-score model, as represented in Table 1 by equation 5 (for individual scores) and by equation 6 (for group mean scores), provides a way by which teachers and experimenters may avoid the worst effects of regression with a minimum of complication. A selection test, denoted A, is used to identify pupils for remedial classes, advanced-placement classes, etc. Then, the pretest, denoted E, is administered just before the treatment is begun. After the treatment, a posttest, denoted B, is administered. Naturally tests A, B, and E must yield equivalent scores. Change is represented by the difference between scores on tests B and E. The fact that selection of pupils is done on the basis of scores on test A and not at all on the basis of scores on test E avoids spurious gains or losses owing to the effect of regression alone.

Table 2 shows that the standard errors of measurement are the same for change scores yielded by equations 1 and 5 as those yielded by equations 3 and 6. It is clear that spurious inflation of change scores yielded by equations 1 and 3 may well cause these changes to be regarded as statistically significant because the standard errors of measurement are the same for both estimates of change. The writer strongly recommends use of equations 5 and 6 whenever a third test can be employed and when maximum simplicity of data analysis is desired.

Another statistical model for making regressed-score estimates of change is represented by equations 7 and 8 in Table 1. The rationale for this model is that the best possible estimate of an examinee's or a group's true score on any fallible test is the estimated true score. A difference between estimated true scores is provided by equations 7 and 8. Note that these equations do not require knowledge of the correlation between pretest and posttest scores. Only the reliability coefficients of these scores are required and, for the most practical purposes, publishers' reliability coefficients may be used after adjustment for range of talent (Davis, 1964, p. 243), which can be accom-

Table 2

Practical Examples

*Basic Data for Davis Reading Test, Series V, in Grade VB, Shadyside School*

All Pupils (N=200)

| Form | Date Administered | Mean | SD | Reliability Coefficient | Sta. Error Measurement | Inter-correlation |
|---|---|---|---|---|---|---|
| A | Sept. 9 | 70.30 | 7.30 | .84 | 2.90 | |
| B | June 4 | 74.50 | 8.00 | .87 | 2.90 | .65 |
| D | June 5 | 74.82 | 8.00 | .87 | 2.90 | |
| E | Sept. 10 | 70.30 | 7.30 | .84 | 2.90 | |
| F | Sept. 11 | 70.34 | 7.30 | .84 | 2.90 | |

| | | Mean Score of 32 Lowest-Scoring Pupils | Hazel | Frank |
|---|---|---|---|---|
| A | Sept. 9 | 59.20 | 60 | 87 |
| B | June 4 | 62.34 | 65 | 92 |
| D | June 5 | 62.66 | 65 | 92 |
| E | Sept. 10 | 60.98 | 62 | 84 |
| F | Sept. 11 | 61.02 | 60 | 87 |

| Equation | Mean Score of Lowest-Scoring 32 Pupils (Mean Change Score) | Hazel (Change Score) | Frank (Change Score) | SE Meas |
|---|---|---|---|---|
| 1. | | 65–60=5.00 | 92–87=5.00 | 4.10 |
| 2. | | 65–61=4.00 | 92–85.50=6.50 | 2.90 |
| 3. | 62.34–59.20=3.14 | | | .72 |
| 4. | 62.50–61.00=1.50 | | | .51 |
| 5. | | 65–62=3.00 | 92–84=8.00 | 4.10 |
| 6. | 62.34–60.98=1.36 | | | .72 |
| 7. | | 56.55–50.40–1.37=4.78 | 80.04–73.08–1.37=5.59 | 3.51 |
| 8. | 54.25–49.73–1.37=3.14 | | | .62 |
| 9. | | 39.65–33.60–1.88=4.17 | 56.12–48.72–1.88=5.52 | 2.40 |
| 10. | 38.03–33.15–1.88=3.00 | | | .42 |
| 11. | | 65.00–67.19=–2.19 | 92.00–86.36=5.64 | 3.56 |
| 12. | | 65.00–65.75=– .75 | 92.00–88.70=3.30 | 3.81 |

plished if the standard deviations of the scores used by the publisher and the scores obtained on the pretest and the posttest are available.

To obtain an estimated true change score from a knowledge of an individual's otbained change score by the usual regression equation (Davis, 1964, equation 2.14), the reliability coefficient and the *variance of the change scores must be computed.* This can be done as follows:

$$r_{CC'} = \frac{s^2_B r_{BB'} + s^2_A r_{AA'} - 2 s_A s_B r_{AB}}{s^2_B + s^2_A - 2 s_A s_B r_{AB}}$$

and $$s^2_C = s^2_B + s^2_A - 2 s_A s_B r_{AB}$$

Then: $$\hat{C} = \bar{C} + r_{CC'} (C - \bar{C})$$

and $$s_{\text{meas } \hat{C}} = r_{CC'} \sqrt{s^2_{\text{meas } B} + s^2_{\text{meas } A}}$$

The last two equations are not given in Table 1 because they require as much data as equation 9 but are less precise.

A third statistical model is represented by equations 9 and 10, which are almost identical with the procedure first presented by Lord (1956). These provide estimates of true change for individuals or for nonrandom subsamples of a group. Essentially, the conventional linear multiple-regression technique is employed to estimate true change from a knowledge of pretest and posttest scores and their correlation. Routines for computing the required ray-score partial multiple-regression coefficients (denoted Wa and Wb in Table 1) have been provided, employing only familiar basic statistics. The standard deviations of pretest and posttest scores denoted (sA and sB) are computed separately because they often differ. The standard errors of measurement for both equations 9 and 10 are given in Table 1 in convenient forms that permit satisfactory tests of the null hypothesis that the true gain is zero.

It should be pointed out that equation 10 is designed for use with a nonrandom subsample of the entire group tested. This equation is not intended for use with nonoverlapping samples drawn successively at random from the same population. If equation 10 is used to estimate the mean change in the entire sample to which a pretest and posttest have been given, the estimate of true mean change, denoted Cs in Table 1, will be equal to B – A, the difference between raw-score means on the posttest and pretest. In the writer's judgment, equations 9 and 10 have greater practical merit than other procedures for estimating change in circumstances under which their use is appropriate and possible. Foresight and planning to make their use feasible in studies involving change are strongly recommended.

Finally, we come to equations 11 and 12, which provide estimates of residual change, which are estimates of deviations from the expected scores of individuals or groups. A zero residual gain indicates that an individual has gained exactly as much as it was predicted that he would, knowing his pretest score and the correlation of the same pretest and posttest scores in an age or grade group of which the individual is a natural member. This is very different from the meaning of a zero change score from equations 1–10; for all of these, a change score of zero indicates that an individual or the mean of a nonrandom subsample of a group was not affected by the treatment provided between the pretest and posttest with respect to the variable measured by those tests. This is what common sense indicates that a zero change score should mean. If it were not for the fact that equation 9 is extremely appropriate for measuring under-and overachievement, the residual-gain techniquest might be employed for that. But as matters stand, they do not seem to have special merit for any purpose.

Equation 11 shows the conventional estimate of deviation from predicted change in terms of raw scores; equation 12 shows a variant of this procedure that causes the change scores to correlate zero with true pretest scores instead of with obtained pretest scores (which is done by equation 11). A comparison of the standard errors of measurement for equations 11 and 12 indicate that the former yields the more accurate scores. Equation 12 is expressed in terms of raw scores instead of in the metric originally used by the authors (Tucker, Damarin, & Messick, 1966).

## REFERENCES

Bereiter, C. Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison, Wis.: University of Wisconsin Press, 1963.

Davis, F. B. The assessment of change. In E. P. Bliesmer & A. J. Kingston (Eds.), *Phases of college and other adult reading programs*. National Reading Conference 10th Yearbook. Milwaukee: the Conference, 1961. Pp. 86-99.

Davis, F. B. *Educational measurements and their interpretation*. Belmont, California: Wadsworth, 1964.

DuBois, P. H., & Manning, W. H. *Methods of research in technical training*. St. Louis: Washington University, 1961.

Lord, F. M. The measurement of growth. *Educational and Psychological Measurement*, 1956, *16*, 421-437.

Lord, F. M. Further problems in the measurement of growth. *Educational and Psychological Measurement*, 1958, *18*, 437-454.

McNemar, Q. On growth measurement. *Educational and Psychological Measurement*, 1958, *18*, 47-55.

Tucker, L. R., Damarin, F., & Messick, S. A basefree measure of change. *Psychometrika*, 1966, *31*, 457-473.

Cronbach, L. J., & Furby, L. How should we measure change—or should we? *Psychological Bulletin*, 1970, *74*, 68-80.

JOHN FOLLMAN
A. J. LOWE
EDWARD UPRICHARD
LOREN ROBERTS
MELVIN VILLEME
University of South Florida

# TYPEFACE AND PLEASINGNESS, LEGIBILITY, PREFERENCE, READING COMPREHENSION AND READING RATE

The objective of this study was to determine the effects of three typefaces, ELITE, PICA, and SCRIPT, on student judgments of pleasingness, legibility, preference, and reading comprehension and reading rate.

There has been considerable research of the effects of typeface variables on legibility, reading comprehension and rate. However, most of this research was concerned with the effects of print, not type, as Payne (1967) noted. Spache (1966) generalized that typesizes between 9 and 12 points can be read equally fast. Tinker (1963) concluded that typefaces then in common use were equally legible. Typefaces most commonly used currently in higher education are probably elite and pica. Campbell (1969) contended that elite has become increasingly preferred to pica. Poulton and Brown (1968) in England examined empirically the effects of upper and lower case typeface combinations of IBM 72 PICA, Smith-Corona 66 ELITE, and Dyeline Siemen typefaces. Small differences in comprehension were obtained, although ELITE texts were easier to read than some other combinations. Another typestyle frequently used in higher education is script. Bell ([illegible]) found that cursive script was read significantly faster than manuscript. There has been little reported research of the effects of currently used typefaces on college students' judgments of pleasingness, legibility and preference or on their reading comprehension or rate.

Recently Follman, Lowe, and Miller (1970) investigated the effects of typeface and also the effects of multiple choice item option format arrangements on college students' comprehension. Four typefaces, ELITE, PICA, PROPORTIONAL, and SCRIPT, each with two multiple choice format arrangements, were used. No significant differences in comprehension were found for typefaces. Since high reliability estimates were found for six of the eight experimental groups (all but the two PROPORTIONAL groups) it was concluded that kind of typeface does not influence comprehension.

The objective of the current study was partly to replicate and partly to extend the above Follman, *et. al.* study. Replicated were the ELITE, PICA, and SCRIPT typeface comprehension conditions. Extensions included additions of students' judgments of typeface pleasingness, legibility, preference, and also reading rate.

## PROCEDURE

The subjects (Ss) were 86 freshmen in an elective introductory education course at the University of South Florida in November, 1971. Ten were males and 76 were females.

Three typefaces were used, ELITE, PICA, and SCRIPT. Each typeface is mediated through a separate typing element. All three typefaces were typed on an IBM Selectric typewriter. ELITE is IBM PRESTIGE ELITE 72 Code 012, a 12 pitch weighted type. ELITE is a relatively small type with 12 characters to the inch. PICA is IBM COURIER PICA 72 Code 015, a 10 pitch square-serif pica type. PICA is a larger type than ELITE but still has 12 characters to the inch. SCRIPT is IBM SCRIPT Code 090, a 12 pitch type which simulates handwriting. SCRIPT is the same size as ELITE with 12 characters to the inch. Thus all three typefaces have the same number of charactesr per inch. ELITE and SCRIPT upper case letters respectively are 2.5 millimeters and 3 millimeters high and 1.5 and 1.5 millimeters wide. PICA upper case letters are 2.5 millimeters high and 2 millimeters wide. The main distinction between ELITE, PICA and SCRIPT is width of letters. Another distinction is that the script characters are linked as in handwriting and are taller.

The IBM Selectric typewriter used has a 12 pitch type which has 12 spaces per inch and is appropriate for the 12 pitch ELITE and the 12 picth SCRIPT typing elements. However, the 10 pitch PICA element has larger letters and is intended for a 10 pitch IBM typewriter with 10 spaces per inch. When the 10 pitch PICA element is used on a 12 pitch machine, which has two more characters per inch than the 10 pitch machine, the effect is to crowd the wider PICA letters together horizontally.

The Pleasingness, Legibility, and Preference ranking materials as well as the reading comprehension and rate materials were typed on stencils and reproduced on a conventional mimeography machine adjusted for 20 pound paper and set for medium slow speed. Spacings, margins, and leading were held constant for each typeface for all conditions. An initial student information sheet containing general information about the experiment was typed in a fourth typeface, a large pica style with 10 characters to the inch.

Three passages with similar content were used for the Pleasingness, Legibility, and Preference rankings. Each passage consisted of about 100 words. The passages were from the *Davis Reading Test Form 1A* (Davis and Davis, 1957).

All six possible typeface orders were used for the three typefaces for the Ss' Pleasingness, Legibility, and Preference rankings. The Ss ranked each set of three typefaces for Pleasingness, Legibility, and Preference independently, using rankings 3, 2, or 1.

Following the Pleasingness, Legibility, and Preference rankings, three typeface comprehension conditions were used, one for each typeface, ELITE, PICA, and SCRIPT. The Ss were randomly assigned to one of these three conditions. *The Nelson-Denny Reading Test Form 1A* comprehension passages and corresponding multiple choice questions were used for the three comprehension conditions. The same passage was typed in each of the three typefaces, and a different typeface passage was administered to each of the three groups. Testing instructions similar to the standard directions were used. Rate was tested first. The Ss read the passage for one minute. At the end of the minute they stopped reading and wrote the line number indicating the number of lines each had read The Ss then resumed reading and answered as many of the questions as they could in the remaining 19 minutes Thirty six items were used for comprehension.

Friedman's analysis of ranks was conducted for the Ss' rankings for Pleasingness, Legibility, and Preference, independently, to determine if significant differences obtained in the Ss' rankings.

Split half and Kuder Richardson 20 reliability estimates were obtained for the ELITE, PICA, and SCRIPT comprehension passages. Simple ANOVAs were conducted to determine if significant differences obtained for typeface for reading comprehension, and reading rate. Raw scores were used for the reading comparisons. For rate the line number was converted to the number of actual words read and the number of words read became the rate measure.

## RESULTS

The Friedman's Analysis of Ranks ($x_r^2$) for Pleasingness indicated a $x_r^2$ of 35.3, significant at the .01 level. These significant differences reflect primarily the 54 of 78 highest rankings awards to SCRIP compared to ELITE and PICA. There were a few more second rankings for ELITE than for PICA. Kendall's Coefficient of Concordance (W) was .23 indicating litle ranking agreement across the three ranks by all 778 rankers.

The $x^2$ for Legibility was 34.5, significant at the .01 level. These significant differences reflected a clear first ranking (43 firsts) for ELITE, a clear third ranking (47 thirds) for SCRIPT, and a ranking of PICA in between. W was .22, again indicating little consistent ranking agreement across all three ranks by all 78 rankers.

The $x_r^2$ for Preference, 2.3, was not significant. ELITE was non-significantly ranked first by 40 students and SCRIPT was ranked first by only 15 students. PICA had similar numbers of ranks of 3, 2, and 1 which apparently prevented signicant findings. W was .02 indicating no consistent ranking agreement across all three ranks by all of the 78 rankers.

For these three sets of rankings the Ns were made equal in each of the six orders, 13 each, for a total of 78. Seventy six of the Ss were females with two males added to balance the six orders.

It is possible that the sample composition of 76 (of 86) females may have introduced a sex bias which may account for some or all of the findings especially the SCRIPT Pleasingness results.

The reading comprehension analyses were run for females only, and also for females and males combined. Table 1 indicates group sample sizes, medians, means, standard deviations, and odd even and Kuder Richardson 20 reliability estimates for females, and females and males combined, for comprehension, for each of the three typefaces, ELITE, PICA, and SCRIPT.

Table 1
Comprehension Test Results

| | N | Med. | $\overline{X}$ | SD | OE | KR |
|---|---|---|---|---|---|---|
| ELITE Females | 21 | 25 | 24.29 | 4.43 | .62 | .79 |
| ELITE Combined | 25 | 22 | 23.40 | 4.59 | .69 | .78 |
| PICA Females | 30 | 20 | 22.40 | 5.68 | .91 | .87 |
| PICA Combined | 33 | 21 | 22.52 | 5.43 | .90 | .86 |
| SCRIPT Females | 25 | 20 | 21.56 | 5.26 | .88 | .85 |
| SCRIPT Combined | 28 | 21 | 22.08 | 5.44 | .88 | .86 |

It is apparent from Table 1 that the inclusion of the males produces little differences in most of the indices. The reliability estimates are all high except for the two uncorrected odd even ELITE groups which are moderately high. All reliability estimates are high enough to justify examination of typeface differences particularly since these estimates are based on a total of only 36 items.

An ANOVA for typeface differences on comprehension for females only indicated a non-significant F of 1.615 for 2 and 73 degrees of freedom. The ANOVA for the combined sample also indicated a non-significant F, .4437, for 2 and 83 degrees of freedom. These findings indicate that typeface, at least the three typefaces used herein under the conditions employed herein, has little if any effect on reading comprehension.

Reading rate means and standard deviations respectively were 353.14 and 89.38 for ELITE, 314.37 and 77.20 for PICA, and 321.40 and 63.74 for SCRIPT, for the female sample. For the combined sample rate means and standard deviations respectively were 344.92 and 84.37 for ELITE, 326.97 and 89.68 for PICA, and 327.39 and 63.54 for SCRIPT. Group sizes were the same as in the comprehension analysis.

The ANOVA for rate for females only indicated a non-significant F of 1.6221 for 2 and 73 degrees of freedom. The ANOVA for the combined sample likewise indicated a non-significant F, .4172, for 2 and 83 degrees of freedom. These results indicate that ELITE, PICA, and SCRIPT typefaces, under the conditions used herein, little influence reading rate.

## CONCLUSIONS

1. University students manifest different rankings to different typefaces on different style and size variables.
2. SCRIPT was clearly more pleasing than ELITE or PICA.
3. ELITE was seen as more legible than PICA or SCRIPT.
4. Typeface has little if any effect on reading comprehension.
5. Typeface has little if any effect on reading rate.

Finally a caveat must be noted. The findings obtained herein reflect the results of the experimental conditions and the characteristics of the sample employed. Since most of the Ss were females a sex bias may have occurred. Therefore replication of this study should be conducted with a larger sample including an equivalent number of males.

## REFERENCES

Bell, H. M. The comparative legibility of typewriting, manuscript, and cursive script: easy prose, letters, and syllables. *Journal of Psychology*, 1939, *8*, 295-309.

Campbell, W. G. *Form and style in thesis writing.* (3rd ed.) Boston: Houghton Mifflin, 1969.

Davis, F. B., & Davis, C. C. *Davis Reading Test Form 1A*. New York: The Psychological Corporation, 1957.

Follman, J., Lowe, A. J., & Miller, W. Typeface and multiple choice option format. In G. B. Schide & M. M. May (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook. (Vol. 1) Milwaukee: the Conference, 1970. Pp. 135-140.

Nelson, M. J., & Denny, E. C. *The Nelson-Denny Reading Test*. Boston: Houghton Mifflin, 1960.

Payne, D. E. Readability of typewritten material: proportional versus standard spacing. *The Journal of Typographic Research*, 1967, *1*, 125-136.

Poulton, E. C., & Brown, C. H. Rate of comprehension of an existing teleprinter output and of possible alternatives. *Journal of Applied Psychology*, 1968, *52*, 16-21.

JAMES L. WARDROP
University of Illinois

# MEASURING CHANGE: SOME PROBLEMS AND SOLUTIONS

Suppose we begin with what *seems* to be a very simple problem: You are responsible for instruction in reading for all third-grade pupils in a school. Notice that this does not mean you are responsible for *instructing* all these pupils. Rather, the job is one of selecting materials and pedagogical approaches, providing teachers with any in-service training they need in order to do an effective job, and developing some sort of pupil motivation system. Furthermore (and here is the catch!), you are to demonstrate at the end of the school year that this group of third-grade pupils has made satisfactory progress in reading skill development during the year. (If you need some further motivation, assume that your job for next year depends on how well you have accomplished this objective. Better yet, assume also that one-third of your salary during this year will be based on the extent to which pupils "make satisfactory progress.")

I am going to presume that you know enough about how to teach reading to third-graders to be able to design and implement a reasonable instructional package. The problem now becomes one of determining what might constitute "satisfactory progress in reading skill development" and of deciding how best to measure that progress. In the following paragraphs, I am going to present two alternatives for dealing with the problem, then reinforce the choice you will have made between the two.

### AN ALTERNATIVE INCORPORATING STANDARDIZED TESTS

Understanding the rationale for the first alternative requires some knowledge of how standardized tests are typically developed. For our purpose, we will consider the process of developing a third-grade level standardized test in reading. Such tests, with minor variations, consist of two parts: a vocabulary subtest and a subtest which is called comprehension. (Within each subtest, further subdivisions may be made, but scores are most often reported only in terms of the two general areas.) Now, a test developer or development team will begin by surveying the most used texts and other curriculum materials in reading at the third-grade level and developing a collection of the kinds of items that might be included in the test: such things as commonly encountered vocabulary words and kinds of reading passages pupils are expected to comprehend. Additionally, the developers consider statements of objectives which have been prepared for the various sets of instructional materials and, generally, the scope and sequence of skills to be taught. After this compilation process has been completed, the job of synthesizing, abstracting, and delimiting is undertaken.

Eventually, (and *that* word subsumes an enormous amount of work), everything is distilled down to a remarkably small number of test items (say, 50 vocabulary words and maybe 40 items measuring comprehension) which represent a remarkably broad range of materials and objectives.

#### *Specification of Skills*

The kind of test development process just described represents one means of defining the appropriate set of reading skills to be developed in grade three. After all, the objectives and skills which survive the selection process and are incorporated into the final version of the test represent an inferred consensus as to what is most important. This consensus, remember, is derived from those instructional materials most widely used throughout this country. We shall return to this issue shortly.

#### *Defining Satisfactory Progress*

Accepting for the moment this definition of the appropriate domain of skills, we are then faced with the question of what constitutes satisfactory progress. Here again standardized tests might provide an answer. After all, most standardized tests have tables by means of which it is possible to convert raw scores into grade-equivalent scores. What could be a more reasonable definition of satisfactory progress than a gain of one year in grade-equivalent scores on some standardized test? (This question is meant to be rhetorical for now.)

With this background, then, we can propose a strategy for determining whether pupils have made satisfactory progress in reading skill development. The strategy goes something like this: about one month into the school year, give all pupils in grade three the appropriate level of some standardized reading test or the reading portion of some standardized achievement test battery. Near the end of the school year, administer an alternate form of that same test. An average gain (in terms of grade equivalent scores) of 0.8 years is now our definition of satisfactory progress. Perhaps we would even choose to deal with the question in terms of individual pupil gains rather than group averages.

*Limitations of this Alternative*

The case for using standardized tests both to represent and to evaluate progress towards skills development objectives certainly seems reasonable. Yet this approach—which has predominated for a good many years—has been criticized on a number of grounds. Some of those criticisms, and references to sources where they have been discussed, are enumerated here. For further details, consult the references cited.

> In an attempt to make standardized tests as representative as possible of *everyone's* objectives and curricula, the test developers have made them an adequate representation of *no one* set of objectives and curriculum (Stake, 1971).
>
> On a typical standardized test of vocabulary, to take just one example, a gain of 0.8 years (grade equivalent) represents at best improvement by four or five items and at worst improvement by one item from fall to spring testing (Feldmesser, 1971).
>
> Over relatively short time intervals (within a single school year, especially), "change" scores based on the difference in performance on a standardized test at two different times are very unreliable. Many, many references could be cited here; among the more recent are Feldmesser (1971), Stake & Wardrop (1971), and Wardrop (1971).
>
> If progress is to be judged on the basis of standardized achievement test performance, the temptation will always be to "teach the test," thereby invalidating it as an indicator of any generalized growth in reading skills (Wardrop, 1971).

These four objections are, separately and collectively, sufficient to warrant the conclusion that sole—or even primary—reliance on standardized achievement tests to measure change in performance in conjunction with a *particular* instructional program in a *particular* location is inappropriate. Standardized achievement tests were never meant to be used as devices for measuring change in this way. This is not to say that standardized tests do not provide useful information. Far from it. The point is that there has in the past been far too much reliance on standardized test scores as serving all the needs for achievement measurement in the schools.

Such a conclusion is a little disheartening, because any other solution to the problem of "satisfactory progress in reading skill development" is going to require much more effort to bring about. Various attempts have been made to modify standardized tests or testing procedures to overcome one or more of the objections cited above. Unfortunately, these have so far met with little success, and some—for example, the efforts in connection with the Office of Economic Opportunity experiment in performance contracting—seem to have been out-and-out disasters.

### AN ALTERNATIVE INCORPORATING JUDGMENTS AND STANDARDS

The second alternative for dealing with the problem has a somewhat different basis from the first. We begin with the notion that an individual's competency in a subject-matter area can be thought of as varying along a continuum ranging from complete ignorance to perfect proficiency. In reading, for example, the inability to make any discriminations involving letters or words represents the lower end of this continuum, while reading with ease such things as *New York Times* editorials or articles in the *Encyclopaedia Britannica* might be reasonable representations of the upper extreme. (It has been suggested that being able to comprehend the instructions for filling out a Federal income tax return represents the ultimate task.) Whether these particular tasks are acceptable representations of these reference points is irrelevant. What is important is that, given a little time for discussion, we could arrive at some mutually acceptable indices of the lower and upper extremes of the scale. Given some additional time for discussion and negotiation, we could develop some tasks representing intermediate points along the continuum. A newspaper comic strip like *Blondie,* a cooking recipe, and the owner's manual for a new car are illustrative of the kinds of reading material which could readily be used to develop some of the kinds of tasks I have in mind.

### *Specification of Skills*

Take this notion of tasks, along with the idea that such tasks represent various points along the continuum from ignorance to mastery, and an interesting strategy for attacking our problem is suggested. Surely we could get some agreement about two or three or six tasks which a pupil who has successfully completed third grade really ought to be able to do. We could probably agree on another several tasks that a pupil who has completed second grade—or one who has completed fourth grade—should be able to perform. Even if we cannot agree on the exact grade placement of tasks, we can at the very least agree as to which tasks should be included. The statisticians and psychometricians—people like me—can empirically place the tasks at the appropriate level based on actual responses by pupils in the schools.

There are a couple of important questions yet to be answered: how do we evaluate pupil performance on these tasks, and how do we use the results as indices of change? If I deal with the easier question first, maybe I will be forgiven for not having a better answer to the hard one. The most direct and useful method of using the results is also one of the simplest; it goes something like this.

### *Determining Satisfactory Progress*

Have pupils perform several of these tasks at the beginning of third grade, and for each task report simply the percentage of pupils who perform it satisfactorily. At the end of third grade, give those same tasks and again report the percentage of pupils performing each task satisfactorily. To illustrate, suppose we have seven tasks, with results as in Table 1.

Table 1
Hypothetical Results for Seven "Reading Performance Tasks" in Grade 3

| Task[a] | Percent of Pupils Performing Satisfactorily | |
|---|---|---|
| | Fall | Spring |
| A | 67 | 95 |
| B | 73 | 92 |
| C | 41 | 88 |
| D | 32 | 73 |
| E | 30 | 80 |
| F | 23 | 60 |
| G | 11 | 42 |

[a] Tasks A and B are "second-grade" level. Tasks C through E are "third-grade" level. Tasks F and G are "fourth-grade" level.

Tasks A and B are "end-of-second-grade" tasks; C, D, and E are "end-of-third-grade" tasks; and F and G are "end-of-fourth-grade" tasks. This information allows us to make some general kinds of judgments about the group. On second-grade tasks, they gained an average of nearly twenty-five percent and were performing with a success rate well above ninety percent. On these tasks which are most relevant—those appropriate to the end of the third grade—the average gain was more than forty-five percent, from an initial level of about thirty-five percent success to a final success rate of about eighty percent.

Are these gains significant? If you ned a statistical analysis to answer that question, the results are not important enough to warrant the effort. If the changes are not large enough to be obvious merely by inspection, they are not likely to be of any practical importance, no matter what a statistical significance test might indicate.

*Some Unresolved Issues*

I have avoided an important aspect of this procedure: the determination of whether a pupil can perform a task successfully. I would *like* to give a solution which made possible the efficient, accurate, economical evaluational of large numbers of pupils, as is possible with standardized achievement tests. Indeed, at one level it might be possible to obtain the desired information through some creative use of traditional kinds of machine-scorable answer sheets. It seems to me that one of the most useful techniques involves the use of reading specialists and performance tasks in combination. If we want to know whether a child can read and "comprehend" one of Aesop's fables, we have him read it in the presence of such a specialist and have the specialist ask whatever questions seem necessary to determine the level of understanding the child has attained. Several such interactions between the reading specialist and the child should make it possible for the specialist to render a judgment about the child's level of reading skill development. Such a procedure is obviously not suited to large-scale evaluations of individual pupils. Neither is it a reasonable substitute for traditional standardized testing programs. Its implementation would, however, create a fantastic job market for those individuals whose training and experience qualify them as reading specialists.

More realistically, we can relatively easily—and the emphasis must be on the relatively—devise easily-administered, easily-scored paper-and-pencil measures based on the notion of performance tasks. Such measures would, in terms of their content, probably look much like current standardized tests. The major difference would be that such measures would contain relatively high proportions of very easy and very difficult items, because an item's inclusion or exclusion would depend not upon its psychometric properties but upon its logical relationship to the underlying continuum postulated earlier. This issue has been discussed more extensively by Wardrop and Stake (1971).

If my presentation of this second alternative for measuring change seems incomplete to you, be assured that the fault is mine, not yours. I am convinced that it represents a more viable strategy than does the current reliance on standardized achievement tests. I am especially convinced that much more thought and even more work will be needed before we have enough of the answers to judge its usefulness. Hopefully, the answers will be found before the questions are forgotten.

## REFERENCES

Feldmesser, R. A. Measurement problems in performance contracting. *NCME Measurement News*, 1971, *14*, 6-7.

Stake, R. E. Testing hazards in performance contracting. *Phi Delta appan*, 1971, *52*, 583-589.

Stake, R. E., & Wardrop, J. L. Gain score errors in performance contracting. *Reading Research Quarterly*, 1971, *6*, 323-325. Also in *Research in the Teaching of English*, 1971, *5*, 226-229.

Wardrop, J. L. Providence to New Century to change: Double play or double cross? Educational Product Report, 1971.

Wardrop, J. L., & Stake, R. E. The inappropriateness of traditional test development procedures for measures designed to assess group differences. Paper presented at meeting of the Psychometric Society, April 1971.

DAVID M. WARK
DIANA TOSTENRUD
University of Minnesota

GEORGE NELMS
Medtronics, Inc.

# HEART RATE AND READING

One thing can be said about reading teachers. Whether classified as remedial, developmental, elementary, secondary, junior college or adult, they are generally sound of mind and pure of heart. However, that same compliment cannot be granted reading scientists. The kindest thing one can say about reading researchers is that they are all sound of mind. Empirically, they seem to be heartless.

An exhausting, if not exhaustive, search of the reading literature using both specific key words like "cardiovascular", "cardiac", "electrocardiogram" and general category descriptors like "physiological" and "neurological" failed to produce one study dealing with heart rate and reading. (The authors wish to acknowledge the invaluable assistance of Dr. Martin Kling of Rutgers University in the bibliographic search.) There were investigations of the seeing, hearing, speaking, and thinking organs and muscles in reading. Mind was investigated in fuzzy and rigorous ways. But apparently no one had ever gone to the heart of the process.

There were two contemporary sources for an interest in cardiovascular variables in the reading process. The first was a recent paper by Wark (1970). Taking a frankly behavioristic position for developing a dynamic model of reading, he proposed strategies for doing reading research in real time. In the context of discussing a then current model of reading (Manzo, 1970), he pointed to four difficulties with the standard psychometric study of the reading process. One of the problems he mentioned had to do with time. Most traditional reading variables cannot be measured rapidly enough to function in a dynamic description. For example, reading rate is usually measured over a time span of from one to three minutes. Yet in that interval there are tremendous fluctuations of behavior (Wark, 1969). This intra-individual variance is usually lost. Again, it may take anywhere from ten to thirty minutes to administer a comprehension test. Typically we assume that comprehension is constant throughout that period. Such standard measures are totally inadequate for a satisfactory dynamic description of reading.

As a solution, Wark proposed what he called a behaviormetric analysis. He suggested starting with peripheral measures of behavior and moving from the skin in. The methodology would involve using available hardware to make on line, movement-by-movement measures of behavior during reading. Such variables as pupil diameter, sub-vocalization, blink rate, retinal ground color, and eye movements measured by the electric Reading Eye II camera would be relevant. These measures could be intercorrelated to give a dynamic picture of the reading process. Heart rate would of course be one such measure.

The second source of interest in heart rate is a concern for emotional factors in reading. More specifically, there are puzzling clinical problems which seem to be motivational in nature. For example, some students can read a paragraph at Time 1. But in the same book, with the same vocabulary and on the same topic, they can't interpret the next paragraph at Time 2. As another example, some students said to be suffering from "test panic" fail to make sense of a written examination question at Time 1. But they can do so immediately after the test has been turned in at Time 2. In a recent paper, Wark (1971) suggested a model for this treatment challenge. He proposed a clinical procedure using deep muscle relaxation and systematic desensitization (Wolpe, 1969). The interest in behavior therapy with avoidance hierarchies to combat a presumably debilitating state of over-arousal flows from this concern. Physiological variables including heart rate are relevant to this concern.

Thus, in summary, there were two wellsprings for the interest in heart rate. One was methodological, an attempt to define a new way to describe reading. The other was clinical, a desire to understand some common problems of emotional nature involved in the reading process.

## REVIEW OF CARDIOVASCULAR RESEARCH RELEVANT TO THE READING PROCESS

The crucial investigation, around which controversy relevant to reading research turns, is the work by Lacey (1959). He critically reviewed the use of physiological approaches to the evaluation of psychotherapy. Two studies are particularly interesting, since they involve reading materials. Rowland (1936) had subjects read typewritten phrases which recalled previously memorized situations. The higher the independently related "exciting character" of the situations, the higher the increase in physiological measures. Mittleman and Wolff (1939) found definite decrease in finger temperature in a subject reading an address by President Roosevelt. The subject, needless to say, had some very strong negative feeling about F.D.R. The other studies reviewed by Lacey involved a variety of verbal and physical stimuli, none of which would be considered reading tasks. Effects were measured by a large number of physiological responses, frequently including heart rate. The results were generally consistent: there was a direct relationship between the "intensity" of the stimulus and the amount of acceleration. These predictable and direct increases are subsumed nicely by activation theory. The stimulus activates various physiological systems, which show an increase or acceleration in heart rate, breathing, galvanic skin response, etc. In fact, such increases as a result of external stimulation are accepted as the general rule in similar investigations. It would not be remarkable to find, therefore, that heart rate increased as a result of any reading input.

Under some conditions, however, Lacey discovered that there can be a response fractionation. That is, under certain conditions some physiological variables such as electrical skin conductance increased, while in the same experiments heart rate decreased. In general, heart rate behaved as expected (went up), during tasks which Lacey characterized as "environment rejecting." These included doing mental arithmetic, completing sentences, and tolerating pain. He speculated that these tasks involved the avoidance or rejection of task-irrelevant aspects of the environment. Lacey pointed out that nerve fiber connections from pressure sensors in the carotid sinus, stimulated by an increased heart rate have been shown to inhibit cortical sensory and motor activity. Such inhibition would be helpful to a subject doing mental arithmetic, who would do best work concentrating on internal symbolic manipulation. For a subject experiencing severe pain, physiological mechanisms to inhibit environmental and proprioceptive inputs would be very useful. On the other hand, Lacey noted that what he called "intake" tasks are associated with heart rate deceleration. These were experiments in which the subject was told to look at flashing lights and be prepared to describe them, or to listen to unpleasant dramatic readings on a tape recorder. The subject was asked primarily to detect the environment, without any specific task-relevant demands. In those situations, heart rate drops. As an illustration of heart rate deceleration in a straight intake task, Lacey cites a study by Davis (1957). The pattern of high sweat gland activity and low heart rate "is produced in its purest form in male subjects presented with a picture of a nude female [Lacey, 1959, p. 199]."

The problem is: where does reading fall on this dimension from environmental intake to environmental rejection? Some reading is probably just casual visual perusal of print, engaged in to fill time with no specific intention by the reader to retain or use the information. This sort of activity probably falls at the acceptance end of the dimension, and would be associated with heart rate deceleration of some sort. On the other end of the dimension might be difficult activities involving reading to learn, engaged in to find answers to relevant questions raised by important people such as teachers, clients, employers, or by information demands of a job. The more difficult this reading, the more likely the reader would be to exclude or reject irrelevant aspects of the environment during reading and consequently show heart rate increases.

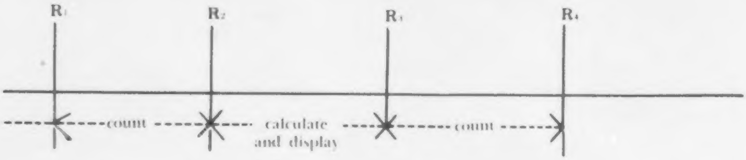
Campos and Johnson (1966) have offered some interesting evidence to clarify the sources of heart rate acceleration. They demonstrate that when a student is instructed that he will have to talk about or vocalize what he has seen, there tends to be an acceleration in heart rate. When the same task is used without instructions about future verbal description, there was a non-significant deceleration in heart rate. Similarly, Weiner (1962) had his subjects looking at TAT cards. He found that in cooperative subjects, the request to make up a story about a card produced heart rate increases. However, he found, quite consistently and reliably, that when subjects were asked to make up a story but told that they did not have to tell it, there were very minimal gains. Buck, Miller and Caul (1969), also found heart deceleration in reading-like intake. He had students look at a series of slides, but without a request to talk about them. Seemingly, requests for future verablization shift reading activity toward the end of the continuum characterized by a rejection of the environment, and are associated with cardiac increase.

Where then do we stand regarding the relationship between heart rate and reading? Readers merely perusing material to kill time, with no specific task demands would probably show some heart rate deceleration. The same readers, acting as experimental subjects in a research project might feel some pressure to perform even though they were explicitly told there would be no questions after reading. They might feel some demands for future questions, and thus show some aspects of a rejection situation. We would expect their heart rates to fluctuate between acceleration and deceleration merely because they were in an experiment. However, subjects presented with the same reading material and told explicitly that they would have to answer questions afterwards should clearly be operating at the rejection end of the dimension, and show some acceleration. In addition, in line with Weiner and Campos and Buck, there should be some acceleration due to verbalizing. Edfeld (1960), McGajgan

(1964), Hardyck (1968), and Zink (1965) all demonstrated that sub-vocalization activities are involved in reading. They indicate that the more difficult the material, the greater the sub-vocalization and presumably, the greater the heart rate acceleration. We might expect then that some readers in a laboratory situation, even without the demand for questions, would show deceleration and while others show acceleration. Subjects reading in a laboratory, with the explicit requirement of comprehension questions would all be expected to show definite signs of acceleration.

### EXPERIMENTAL INVESTIGATION

Volunteer subjects at the University of Minnesota participated in som investigations of heart rate during reading. The methodology was essentially similar for all subjects reported. After being introduced to the experimental room and given time to relax, three zinc cardiotachometer electrodes were attached, two on one arm and one on the other, using commercial cardiotachometer electrode paste. Data was produced by a Medtronic Digital Cardiotachometer of unique design. The machine starts counting at the onset at the R-wave spike R$_1$. It stops counting at the onset of the next R-wave spike R$_2$. (Typically the time between R-waves will vary from 0.5 second to 1 second). The time counted during the interval R$_1$R$_2$ is converted to a frequency which is displayed digitally in terms at beats per minute during the interval R$_2$R$_3$. The next R-wave spike restarts the cycle. In essence, the machine times each heart beat, and computes what the rate would be if the subject's heart were beating constantly at that rate. The numbers projected by the cariotachometer averaged to yield heart rate in each of the experimental conditions.



Subjects were requested to rest with their eyes closed in order to produce base heart rate. They were then presented with cards from the High School, College, and Adult level of the Reading Eye Test (McGraw-Hill E.D.L., 1966). The students were told explicitly whether there would be tests on the reading. The students alternated reading and resting. The resting period mean was considered base rate; the reading period means was considered task period.

The results reported in Table 1 are generally consistent with the derivation from Lacey's formulations. Under the no question or intake condition, there was an overall drop from resting base rate. Only four of the fifteen differences were in the accelerating direction, while twelve were in the decelerating direction. The averages, clearly were in the deceleration or intake direction. (Hypothetically, if a non-obtrusive measure of heart rate could have been taken, all of the differences would be negative. However, it's difficult to think of a really unobtrusive measure of heart rate involving three electrodes and a cardiotachometer.) In the question or rejection condition, fifteen of the sixteen differences were positive, presumably reflecting the effects of both the experimental situation and the increase in subvocalization. According to the Mann-Whitney U test, the hypothesis that there are no differences between the two conditions can be rejected at $P < .001$. Since the scores on the ten item comprehension test were almost maximum for both subjects in the question condition ($S_3X=9.6$; $S_4X=9.3$) there was not sufficient variability to relate comprehension to heart rate acceleration. Future research might be designed to investigate the relative contribution of the experiment effect and the vocalization effect.

Table 1

Mean Differences in Heart Rate for Resting and Reading

| Subjects | Test Card | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | D |
| | No Questions | | | | | | | | |
| $S_1$ | +0.7 | −0.3 | −2.6 | −0.7 | −1.5 | +1.1 | −7.2 | | −1.0 |
| $S_2$ | +5.8 | +7.6 | −1.0 | −4.6 | −8.1 | −9.9 | −6.5 | −13.6 | −3.7 |
| | Questions | | | | | | | | |
| $S_3$ | +2.7 | +3.3 | +9.2 | +3.0 | +3.6 | +1.2 | +2.5 | +5.8 | +4.2 |
| $S_4$ | +6.7 | +0.4 | +0.5 | +4.8 | +3.1 | +6.5 | +5.8 | −0.4 | +4.4 |

$p(Q = NQ) < .001$

### Conclusions

Heart rate seems to be implicated in the reading process in a way not previously suspected. The nature of the relationship is still tentative. There are many theoretical as well as methodological problems to be solved, but the effort appears to be justified. It may shed a new light on the relation between physiological variables and comprehension. It opens up further research in one of the, as yet untapped, aspects of reading.

We might speculate at this point about the relationship between heart rate and comprehension. When the subject knows he is to be tested, his heart rate goes up. It would be fascinating to know the limits of that relation. Suppose that under some conditions, or for some students, the cardiac acceleration goes "too high." Lacey reports that the stimulation of the carotid sinus pressure sensors may inhibit brain function. In the reading situation, there should be at some level a negative correlation between heart rate acceleration and comprehension test score. Too much acceleration produces enough inhibition to lower comprehension. In such students, training in comprehension may be associated with some demonstrable changes in the amount and direction of heart rate change. Alternatively, the equation may be approached from the opposite direction. It would be fascinating to find that there is an increase in comprehension by highly aroused readers trained to lower their own heart rate. Presumably they would thereby become more relaxed and better able to read. The work in operant control of human heart rate by Shapiro, Tursky, and Schwartz (1970) raises that interesting speculation. Perhaps one piece of hardware in the developmental reading laboratory of the future will be a bio-feedback machine. It will be used to teach students to decelerate their own heart rate. Hardyck (1968) successfully used similar equipment to teach students to reduce excessive subvocalization.

Finally, these results cast some interesting light on the notion of mathemagenic behaviors defined by Rothkopf (1966). He found that questions after reading led to better learning. Perhaps the fact that the questions will appear produced heart rate acceleration as a mediating link or as a correlate of better comprehension. It may turn out that certain levels of heart rate are mathemagenic.

This paper opened with two justifications for the study of heart rate in reading; one methodological and one clinical. Obviously the methodology problems have just begun. How they will be solved is an open question. It should be obvious that the clinical suggestions are way ahead of validated techniques. But one thing can be said, the paper offers researchers a healthy, hearty new area.

### REFERENCES

Buck, R., Miller, R. E., & Caul, W. F. Heart rate and skin conductants responding during three attention-direction tasks. *Psychonomic Science*, 1969, *15*, 291-292.
Campos, J. J., & Johnson, H. J. The effects of verbalization instructions and visual attention on heart rate and skin conductants. *Psychophysiology*, 1966, *2*, 305-310.
Davis, R. C. Responses to patterns. *Transactions of the New York Academy of Science*, 1957, *19*, 731-739.

Edfelt, A. *Silent speech and silent reading*. Chicago: University of Chicago Press, 1960.

Hardyck, C. D. The effects of subvocal speech on reading. Final report project #3256. United States Office of Education, 1968.

Lacey, J. I. Psychophysiological approaches to the evaluation of psychotheraputic process and outcome. In E. A. Rubinstein & M. B. Parloff (Eds.), *Research in psychotherapy*. (Vol. 1) Washington, D. C.: American Psychological Association, 1959.

Manzo, A. V. A theoretical model of the teaching-learning factors believed to be interacting in the development of reading comprehension. In G. B. Schick & M. M. May (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook. (Vol. 2) Milwaukee: the Conference, 1970. Pp. 125-127.

McGaigan, F. J. *Silent speech during silent reading*. (CRP-1863) Hollins College, Va.: Hollins College, 1964.

Mittlemann, B., & Wolff, H. G. Affective states and skin temperature: experimental study of subjects with 'cold hands' and Raynaud's syndrome. *Psychosomatic Medicine*, 1939, *1*, 271-292.

The Reading Eye Test Selections. Huntington, N. Y.: McGraw-Hill / EDL, 1958.

Rothkopf, E. Learning from written instructive materials: an exploration of the control of inspection behavior by test-like events. *American Educational Research Journal*, Nov. 1966, *3*, 241-250.

Rowland, L. W. The somatic effects of stimuli graded in respect to their exciting character. *Journal of Experimental Psychology*, 1936, *19*, 547-560.

Shapiro, D., Tursky, B., & Schwartz, G. E. Differentation of heart rate in systolic blood pressure in man by operant conditioning. *Psychosomatic Medicine*, 1970, *32*, 417-423.

Wark, D. M. Designs for measurement of status. In J. A. Figurel (Ed.), *Reading and Readism*. International Reading Association 13th Proceedings. (Part I) Newark, Del.: the Association, 1969. Pp. 494-501.

Wark, D. M. A dynamic theory of reading. In G. B. Schick & M. M. May (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook. (Vol. 1) Milwaukee: the Conference, 1970. Pp. 320-326.

Wark, D. M. Test panic, procrastination and day dreaming. In D. M. Wark (Ed.), *College and adult reading*. North Central Reading Association 6th Yearbook. St. Paul, Minn.: the Association, 1971. Pp. 131-151.

Weiner, H. Some psychological factors related to cardiovascular responses: a logical and empirical analysis. In R. Roessler & N. S. Greenfield (Eds.), *Physiological correlates of psychological disorder*. Madison: University of Wisconsin Press, 1962. Pp. 115-142.

Wolpe, J. *The practice of behavior therapy*. New York: Pergamon Press, 1969.

Zink, B. P. An electromiographic appraisal of inner speech during silent reading of 36 fourth grade pupils with an IQ range of 90 to 110. (Doctoral dissertation, University of Texas) Ann Arbor: University Microfilms, Order No. 65-8100. *Dissertation Abstracts*, Aug. 1965, *26*, 916.

JOHN A. WOOD
ANNE MARIE BERNAZZA HAASE
University of Hartford

# APTITUDE INTERACTION ANALYSIS: A NECESSITY IN READING RESEARCH

For the past two decades, researchers have expended vast quantities of time and effort analyzing the process called reading. We have studied learner characteristics, teacher variables, school factors, method effectiveness and even parental influence. In the process, a staggering list of factors have been identified, measured and correlated. Thousands of hours have been spent attempting to design studies which would measure the effects of these factors. Most of these studies have dealt with measuring the main effects of specific isolated factors, rather than the interactions among combinations of factors. In general, these attempts have yielded little information. The results are a loss in time and effort, wasted information that remains hidden in the data, and little usable information upon which to base future research.

Cronbach (1957) argued that research efforts should "deal with treatments and persons simultaneously. Treatments are characterized by many dimensions; so are persons. The two sets of dimensions together determine a payoff surface . . . we should design treatments, not to fit the average person, but to groups of students with particular aptitude patterns [p. 315]."

Instructional methods as well as educational environments differ and these differences interact with personal variables of the learner to either facilitate or inhibit learning. Given a common set of objectives, some students will be more successful with one instructional program and other students more successful with alternative programs. Individualized programs support the notion that learner characteristics must be considered in setting up individual programs. However, patterns of interaction between personal learner variables and specific educational methods have yet to be determined. As Glaser (1967) pointed out, there is little value in the *general laws* of learning. Insufficient attention has been paid to specific learner characteristics and their effects in specific learning situations. Without this knowledge, attempts at individualized instruction are limited to broad concepts which do little to adjust the educational climate so that it can accommodate specific learner characteristics and maximize learning. Instead of predicting main efforts between treatments, we must begin predicting interactions based on our knowledge of sociological, personalogical and methodological factors that affect learning.

Although research evidence indicates that boys have more difficulty learning to read than girls, there is little research to establish causation. Two possible reasons may be suggested for this. First, many of the studies are descriptive. Second, the existing experimental research has concentrated only on main effects rather than interaction effects.

## PURPOSE

In order to demonstrate the effectiveness of interaction analysis as a research tool, main and interaction effects of three factors of teacher sex within school, treatment and gender of student across pre- and post-test measures were hypothesized. The selection of the three factors is based on research that has identified them as contributors to th variance of reading achievement scores.

### *Factor I: Student Gender*

Since studies such as Parsely (1964), Gates (1961), Anderson, Hughes and Dixon (1957), Huges (1953) on the relationship between sex differences and reading achievement have consistently emphasized the advantage of female students over male students, student sex was identified as a consideration in structuring meaningful learning situations. This is true both at the elementary and secondary levels.

### *Factor II: Teacher Gender*

One frequently stated rationale used to explain the evidence that more boys than girls experience difficulty in learning to read is that women teachers are unable to adjust to boys. Boys are described as being more difficult to motivate, as having different learning styles than girls and as being more aggressive than girls. Peebles (1969) noted that educational theorists suggest that "women teachers . . . treat boys less favorably than girls, and discriminate against them in subtle ways [p. ]." Research also supports this theory. Meyer and Thompson (1956), Davidson and Lang (1960), and Bond and Dykstra (1967) support importance of teacher sex in the educational environment.

### *Factor III: Training Material*

Materials from different content areas present special reading problems for students. The majority of studies support the contention that reading ability is not consistent across tests of different types of materials. Studies such as Pressey and Pressey (1921), Tinker (1932) Shores (1943) and Fay (1953) support the influence of the kinds of instructional materials used.

## PROCEDURE

Two schools participated in the study. In each school, two experimental groups were formed from the total number of seventh grade students of average reading ability. The rationale for using students of average reading ability is supported by previous research. Peebles (1969) studied the effect of teacher sex with "disabled" readers and concluded that teacher-sex was not a significant factor. Anderson, Hughes and Dixon (1957) concluded that high ability seems to contribute to a diminuation of sex differences.

From the total number of seventh grade students in each school, those students were selected whose total reading comprehension scores taken from school records (administered in the fall of 1968) fell between the points 6.0 and 7.0 inclusive. These students were grouped on a male-female basis.

From this restricted sample in each school, students were randomly selected and randomly assigned to treatment groups. Thus in each school two comparable treatment groups were created. Each class contained 10 students (5 male, 5 female). The total number across both schools involved 40 students.

The specific treatment was randomly assigned to the group. An experienced female reading teacher taught both experimental groups in School 1 and an experienced male reading teacher taught both treatment groups in School 2. This resulted in the distribution of school and teacher effects over both treatment groups.

### *Treatment Groups*

Treatment group I was given thirty reading lessons using material selected from three current seventh grade basal reading programs. Treatment group II was given thirty reading lessons using prepared lesson plans based upon material from three widely used seventh grade social studies texts.

### *Hypotheses*

The hypotheses for the main effects under consideration were: There is no significant difference between School 1 and School 2, between Treatment 1 and Treatment 2, between boys and girls on pre-test and post-test measures: and there is no significant difference between pre-test and post-test measures.

The hypotheses for the interactions of the between variables were: There is no significant interaction between schools by treatment (SxT), between schools by gender (SxG), between treatment by gender (TxG), and among schools by treatment by gender (SxTxG).

The hypotheses for the interactions of the within variables are: There is no significant difference between measures by schools (MxS), between measures by treatment (MxT), between measures by gender (MxG), among measures by schools by treatment (MxTxS), among measures by treatment by gender (MxTxG), and among measures by schools by treatment by gender (MxSxTxG).

### *Design*

A 2x2x2x2 mixed factorial design with three variables between and one within was used (Myers, 1966). The three between subject variables were school, gender and treatment, each containing two lvels; the one within variable was pre- and post-test measures obtained from Forms A & B of the *Sequential Tests of Educational Progress,* Reading Test, Junior High level.

## Results

The Analysis of Variance sumary table of the three between and one within subject variables is presented in Table 1. The main effects for the between subject variables and the interactions between these variables were found not significant. The main effect for the within variable (measure) and the accompanying interactions were found significant. An inspection of the means for the main effect of the within subject variable showed the mean post-test measure ($\overline{X} = 17.56$) was significantly different from the mean pre-test measure ($\overline{X}=16.55$). These data indicate that all students performed better on the post-test, regardless of their school, gender, or treatment.

All the interactions on the measure variable were found significant. Figure 1 presents the ordinal interaction of school 1 and 2 on the pre- and post-test measures.

Subjects in school 2 with a male teacher ($\overline{X} = 17.60$) perform better than subjects in school 1 ($\overline{X} = 15.50$) with a female teacher on the post-test, even though no main effect for school was obtained. The reason for this is that since the $S_1$ and $S_2$ effects are in the opposite direction, they tend to cancel out each other.

A disordinal interaction is found in Figure 2, which shows the interaction of two levels of treatment on two levels of measures. The implication of this interaction is that subjects in $T_1$ ($\overline{X} = 18.00$) showed greatest overall learning as measured by the post-test. Subjects in $T_2$ learned little from the pre- to post-tests ($\overline{X}$ Pre $= 16.72$; $\overline{X}$ Post $= 17.12$). The main effect of treatments was non-significant due to the nullifying effect of $T_2$ on $T_1$ on the measure variable.

Again a disordinal interaction is seen in Figure 3 which represents

## Table 1

### Analysis of Variance Summary Table for the Between and Within Subjects Variables

| Source of Variance | df | Sums of Squares | Mean Square | F |
|---|---|---|---|---|
| Between Subjects | 39 | | | |
| Schools (S) | 1 | 5.5125 | 5.5125 | .328 |
| Treatments (T) | 1 | 0.3125 | .3125 | .019 |
| Gender (G) | 1 | 13.6125 | 13.6125 | .810 |
| SxT | 1 | 21.0125 | 21.0125 | .125 |
| SxG | 1 | 2.8125 | 2.8125 | .167 |
| TxG | 1 | 5.5125 | 5.5125 | .328 |
| SxTxG | 1 | 12.0125 | 12.0125 | .715 |
| Error (between) | 32 | 537.6000 | 16.8000 | |
| Within Subjects | 40 | | | |
| Measures (M) | 1 | 71.0125 | 71.0125 | 7.10• |
| MxS | 1 | 124.6125 | 124.6125 | 12.46•• |
| MxT | 1 | 172.8125 | 172.8125 | 17.28•• |
| MxG | 1 | 226.5125 | 226.5125 | 26.65•• |
| MxSxT | 1 | 139.1125 | 139.1125 | 13.91•• |
| MxSxG | 1 | 193.3125 | 193.3125 | 19.33•• |
| MxTxG | 1 | 79.6125 | 79.6125 | 7.96•• |
| MxSxTxG | 1 | 166.4875 | 166.4875 | 16.65•• |
| Error (within) | 32 | 320.0000 | 10.0000 | |
| Total | 79 | | | |

• $P \leq .05$
•• $P \leq .01$

the performance of boys and girls on the test measures. The main effect of gender was not significant due to the constant performance of the females on the pre- and post-tests. It may be said that boys ($\overline{X} = 17.65$) learned more than girls ($\overline{X} = 17.45$) on the post-test.

As the first order interaction effects occur when differences among T, S, and G change significantly over levels of M, second order interaction effects occur when differences among GS, TS, and GT change significantly over levels of M. These second order effects are presented in Figures 4 and 5.

Figure 4 presents the mean performance of School 1 and 2 and Treatment 1 and 2 on pre-and post-test measures. Both ordinal and disordinal interactions occur. In School 1, both Treatment 1 and 2 affect learning positively. In School 2 only Treatment 1 ($\overline{X} = 17.20$) produces gain in learning as measured by the post-test.

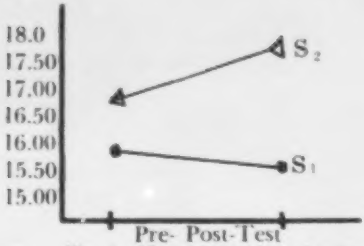


Fig. 1. Mean Scores by schools.

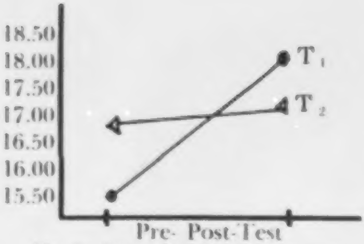


Fig. 2. Mean Scores by Treatments. (M x T)

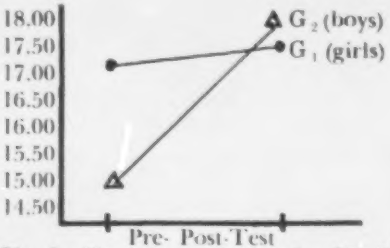


Fig. 3. Mean Scores by Gender.(M x G)

In Figure 5, the interactions among school, gender and measures are shown. Girls in School 1 ($\overline{X} = 18.30$) and boys in School 2 ($\overline{X} = 18.60$) performed better on the post-test than girls ($\overline{X} = 16.00$) in School 2 and boys in School 1 ($\overline{X} = 16.70$).

In the treatment by measure by gender interaction, girls ($\overline{X} = 18.3$) and boys ($\overline{X} = 17.7$) in Treatment 1 and boys ($\overline{X} = 17.6$) in Treatment 2 performed better on the post-test. Girls ($\overline{X} = 16.6$) in Treatment 2 performed very poorly on the post-test as compared to the other three groups and their own pre-test scores ($\overline{X} = 17.7$).

In analyzing the third order interactions, it is well to remember that these effects occur when differences among STG change significantly over levels of M. The third order interactions are presented in Figures 6 and 7. In Figure 6, it is observed in School 1 Treatment 1, both girls ($\overline{X} = 17.60$) and boys ($\overline{X} = 16.8$) improved, with girls superseding boys on the post-test. In Treatment 2 School 1, girls ($\overline{X}$ pre $= 16.0$; $\overline{X}$ post $= 19.0$) improved greatly while boys ($\overline{X}$ pre $= 17.0$; $\overline{X}$ post $= 16.60$) slightly decreased in their performance.

In Figure 7, School 2 Treatment 1 both girls ($\overline{X} = 19.0$) and boys ($\overline{X} = 19.6$) improved. In Treatment 2, a distinct disordinal interaction occurred. Boys increased their performance greatly on the post-test measure ($\overline{X} - 18.60$) while girls dropped significantly ($\overline{X}$ pre $= 19.0$; $\overline{X}$ post $= 14.20$).

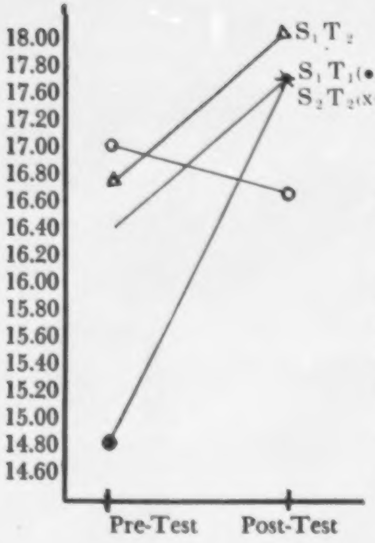


**Fig. 4. Mean Scores by Test Measures Treatments, and Schools. (M x T x S)**

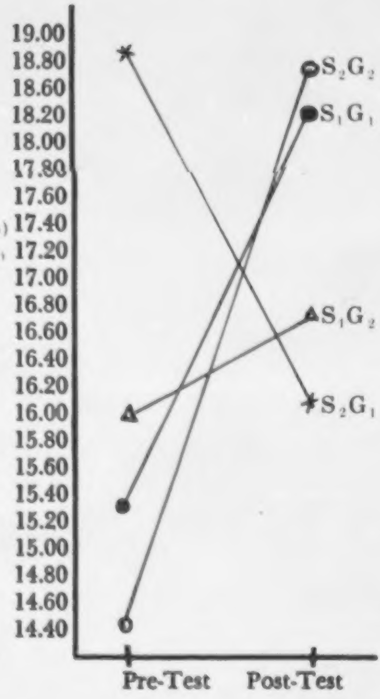


**Fig. 5. Mean Scores by Test Measures, Gender, and Schools. (M x G x S)**

## INTERPRETATIONS AND CONCLUSIONS

This data indicates that a significant gain in reading ability occurred. When this total gain is analyzed by the overall gain in each school, it was discovered that students in School 1 with the female teacher did not achieve as well as students in School 2 with the male teacher. However, any conclusions drawn from these first order interactions without an analysis of the second and third order interactions that occur between the schools, treatments and student sex factors on the within variable measure would be erroneous.

In an analysis of second order interactions, more meaningful information appears. The female teacher in School 1 could work effectively with Treatment 1 but apparently not with Treatment 2. The

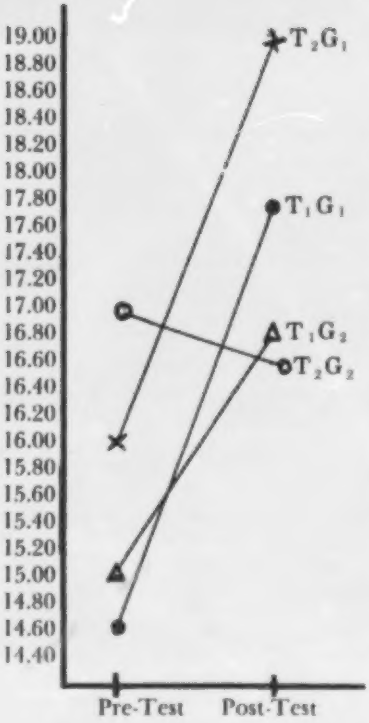


**Fig. 6.** Mean Scores by Test Measures, Treatment, and Gender for School 2 (Female Teacher).

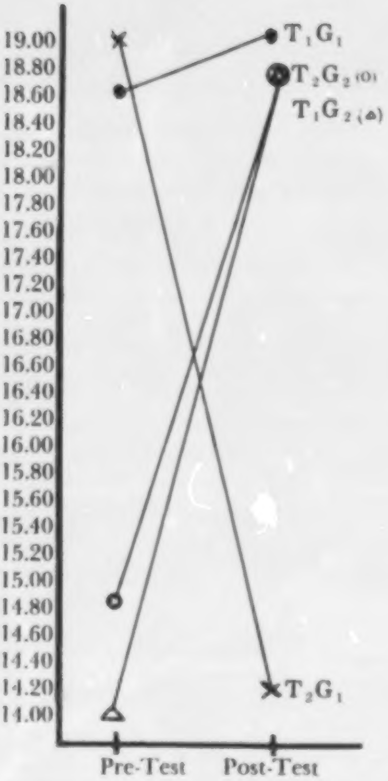


**Fig. 7.** Mean Scores by Test Measures, Treatment, and Gender for School 2 (Male Teacher).

question that remains to be answered is: Is this an accurate interpretation of data or are the effects of teachers compounded by the third order interactions? To answer this question, it was necessary to examine the performance of male and female students within each school.

It can be seen that female students in School 1 with a female teacher performed better than the female students in School 2 with a male teacher. The reverse is true for male students. Male students performed better with the male teacher in School 2 than male students with the female teacher in School 1. Thus the interaction effect of teacher and student sex within school is supported. This also lends support to the theory that male students respond better to male teachers; female students respond better to female teachers. Questions now may be raised concerning the third order intractions among teacher, pupil and treatment.

With a female teacher in School 1 both treatments were effective for female students. However, only Treatment 1 (narrative material) was effective for male students. In School 2 with a male teacher both treatments were effective for male students. However, only Treatment 1 was effective for female students. In other words, with Treatment 2 (text material) the male teacher in School 2 was effective only with male students. When the female teacher in School 1 used Treatment 2, she was only effective with female students.

This study supports the following conclusions.

1. Interaction analysis allows researchers to draw maximum information from the data; to gain insights into interactions at many levels, and to obtain more information upon which to generate new hypotheses.

2. Given the understanding that teacher sex and school effect were confounded, female teachers were more effective with female students than male students and male teachers were more effective with male students than female students.

3. The effectiveness of text material as a training vehicle is dependent upon the interaction between teacher sex (confounded by school effects) and student sex.

## REFERENCES

Anderson, I., Huges, B., & Dixon, R. The rate of reading development and its relation to age of learning to read, sex, and intelligence. *Journal of Educational Research*, Mar. 1957, *50*, 481-494.

Bond, G. L., & Dykstra, R. The cooperative research program in 1st grade reading instruction. *Reading Research Quarterly*, Summer 1967, *2*(4), 5-142.

Cronbach, L. J. The two disciplines of scientific psychology. *American Psychologist*, 1957. *12*, 671-684.

Davidson, H., & Lang, G. Children's perceptions of their teachers' feelings toward them related to self perception, school achievement and behavior. *Journal of Experimental Education*, 1960, *29*, 107-118.

Fay, L. C. How can we develop study skills in the different curriculum areas. *Reading Teacher*, 1953, *6*, 12-18.

Gates, A. I. Sex differences in reading ability. *Elementary School Journal*, May 1961, *61*, 431-434.

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagne (Ed.), *Learning and individual differences*. Columbus: Merrill, 1967.

Huges, B. Sex differences and reading achievement in the early grades. *Clinical studies in reading*. Experimental Education Monograph No. 77. Chicago: University of Chicago Press, 1953. Pp. 102-106.

Myers, J. *Fundamentals of experimental designs*. Boston: Allyn and Bacon, 1966.

Meyer, W. J., & Thompson, G. C. Sex differences in the distribution of teacher approval and disapprovel among 6th grade children. *Journal of Educational Psychology*, 1956, *47*, 385-396.

Parsely, K. M. Further investigation of sex differences in achievement of under-, average, and overachieving students within 5 IQ groups in grades 4 through 8. *Journal of Educational Research*, 1964, *57*, 267-268.

Peebles, J. D. A study of the effects of teacher sex and group composition on the reading progress of disabled boy readers in the third grade. (Doctoral dissertation, Syracuse University) Ann Arbor: University Microfilms, Order No. 70-10,378. *Dissertation Abstracts*, June 1970, *30*, 5326A.

Pressey, L. W., & Pressey, S. L. A critical study of the concept of silent reading ability. *Journal of Educational Psychology*, 1921, *12*, 25-31.

Shores, J. H. Skills related to the ability to read history and science. *Journal of Educational Research*, April 1943, *36*, 534-594.

Tinker, M. A. The relationship of speed to comprehension in reading. *School and Society*, 1932, *36*, 158-160.

IRENE ATHEY
The University of Rochester

# DEVELOPMENTAL PROCESSES AND READING PROCESSES— INVALID INFERENCES FROM THE FORMER TO THE LATTER

Most educators believe that, while teaching is an art, there is a body of established knowledge which constitutes a science of education. The basic content of this science, however, seems to be somewhat difficult to define. One can, of course, make analyses of teacher behavior, or conduct field studies comparing the consequences of organizational, or curricular intervention. Such enterprises constitute the bulk of doctoral dissertations and faculty research emanating from our graduate schools of education. Some sciences, especially those which employ so-called hard data, tend to look down on these kinds of activities as imprecise, lacking objectivity and, in general, barely meriting the label scientific.

People in education have another function or responsibility, which is perhaps no more objective, but certainly just as difficult as the task of instilling rigor into educational research. I am referring to the task of keeping abreast with developments in the sciences (not to mention social and political events which may also have profound effects on education but are not under consideration here) in order to determine how these developments impinge, or should imping, on our educational beliefs and practices. Having determined the relevancy of a particular field to education, the educational expert has the following tasks: a) He must assess the validity and relevance of findings from a particular scientific discipline. Since he is unlikely to be an expert in the field (and is certainly not going to be an expert in all of them), and since experts in the field frequently disagree about these matters, the educator finds himself in a role rather like that of a congressman in committee. He must weigh the various testimonies, and draw his own conclusions. b) Then he must attempt to synthesize, or at least balance the inputs from the different disciplines, to discover whether they are at best congruent, at worst contradictory, but more likely incomplete, piecemeal, and limited in perspective, so that they appear unrelated and incapable of synthesis. c) Next, from the above information he must draw valid interpretations concerning the implications of these findings for educational theory. d) Finally, he must consider the most effective ways to implement the conclusions he has drawn in the practical setting of the schools. The educator's task has been compared to that of the juggler who must keep many balls in the air at the same time, but for the juggler all balls are equal in appearance, weight, and value. The educator has the additional problem of determining which balls to use, and the relative worth of those he chooses.

I believe that the processes of weighing, synthesizing, and interpreting the many and varied inputs we in education receive from the sciences (again, I am omitting, without discounting, the inputs from social and political sources to which education must constantly lend a listening ear), are extraordinarily difficult, and that the recurring phenomenon known as the bandwagon effect may be attributable, at least in part, to this difficulty. Hence we find a tendency among educators to latch on to the findings of some biologist, psychiatrist, or sociologist, and to elevate his pronouncements to the status of a cult, especially if they happen to be in line with their own educational philosophy. A psychologist speaks, and programmed instruction is hailed as an innovation which will change the face of education. A sociologist speaks, and we may well find ourselves deschooling society and issuing educredit cards in the maternity ward.

It is important, of course, that education remain an open system capable of receiving and using data from the sciences, but as the discipline where synthesis and evaluation of the input must also take place, education has the responsibility of maintaining perspective and of having the strength to withstand scientific incursions which are excessive or unduly limited in their perspective. This may sound conservative, and many educators are afraid of being thought reactionary, out-of-date, or apathetic. We should indeed be open, flexible, and innovative, we should welcome an interdisciplinary approach to the problems of education, but we should be on guard against being subject to every wind that blows from every quarter to the academic globe.

My introductory remarks have been couched in terms of education in general but they apply with equal force to reading. For many years, reading has been the exclusive province of the education profession, and its career both in research and practice, has followed the ups-and-downs of education. For example, during an era of progressive education, the tendency was to delay the teaching of reading, and to use introductory methods which placed emphasis on units of meaning, such as the whole word or phrase, whereas the no-nonsense, back-to-basics philosophy (e.g. the Rafferty "Education in Depth" approach) called for early and specific training which would give children the tools—usually phonics training; the research of this era is concerned largely with the efficacy of Method X versus Method Y. Similarly a period which emphasized the education of the "whole child" saw the ascendancy of the maturation hypothesis linking reading age with such physiological measures as skeletal age, dental age, etc. as well as psychological measures such as mental age (Olson, 1949).

Like education itself, reading is today the focus of interdisciplinary interest and effort. Personally, I welcome this expansion, this opening of windows to let more sunshine in. Just let us be careful that, if our position happens to be near one window, we are not blinded by the light from that particular window, so that we cannot see the view from other windows. In this paper, I have attempted to take some of the most recent knowledge from other disciplines notably psychology and linguistics, and to use them, not to put the whole child together again, since I feel this may be premature, but to achieve in some measure the kind of balance I have alluded to above. I will discuss briefly some findings in perception, cognition, psycholinguistics, and finally motivation. In each case, the development of these processes in the non-reading child will be the focus of consideration, rather than the mature processes of the skilled reader.

## PERCEPTION

We may accept at the outset Haber's (1969) view, which is shared by many other researchers, that "sensation, perception, memory, and thought must be considered on a continuum of cognitive activity. They are mutually interdependent and cannot be separated except by arbitrary rules of momentary expediency [p. 1]." It appears that all processes on the continuum are geared to the single objective of reducing the uncertainty experienced by the organism in confronting both his inner and outer world. While recognizing the continuum, psychologists still find the terms perception, cognition, etc. useful.

There are many exciting developments in the field of perception, and we can refer to only a few of them. Infant perception has been studied intensively, and we now have evidence that the infant's world is not the booming, buzzing confusion hypothesized by William James, but is characterized even as early as 16 weeks, by pattern discrimination, object permanence, size constancy, and depth perception. Fantz (1970) and Miranda (1969) at Case Western Reserve University's Perceptual Development Laboratory have found that infants show visual preference for complex, brightly patterned or colored, moving stimuli. In 1961, Walk and Gibson developed the simulation of a "visual cliff" by means of a divided patterned surface. Subsequent research on a variety of species has shown that all except flying and swiming animals avoid the "cliff edge" at a very young age, e.g. at three days of age for monkeys (Rosenblum & Cross, 1963). Infants cannot be tested on this situation until they begin to crawl, but Walk (1966) found that 90% of infants between 6 and 16 months make avoidance responses.

Other areas of study with infants are the perception of size constancy over distance, linking of visual and tactual cues in perception of objects, and the relationship between the position of a stationary object and its movement from place to place. All confirm the presence of stability and coherence in the infant's perceptual world. In fact, Denis-Prinzhorn (1961) found that after infancy, there is even a trend toward over-constancy in size judgments.

Bower (1966, 1971) has conducted some ingenious experiments which show that the very young child is capable of more and finer discriminations than previously suspected, and can register most of the visual information adults can, though they are able to process and use lesser amounts.

If we now consider the data on early infant perception in conjunction with Gibson's (1969) work on perception of letters by means of their distinctive features, the conclusion might be irresistible that teachers could take advantage of this early development to begin teaching recognition of letters, word shapes etc. at a much younger age than is usually the case. Proponents of the academic preschool, for example, might see these studies as presenting yet further evidence that we are failing to capitalize on the young child's abilities and are wasting valuable years of potential learning in 'meaningless play.' Such a conclusion however would represent an inference from a limited segment of data. For one thing, we do not know precisely the operations of the perceptual mechanisms in the adult reader, much less how well such operations are matched in the perception of the young child. Most theorists in this area see the organism as inhabiting a world of 'noise' from which he needs to extract such information as will reduce his uncertainty about present and future events, but they disagree as to the mechanisms which are used to this end. Moray (1969) includes the following elements: mental concentration, vigilance (paying attention in the hope the event will occur), selective attention (selection of one of several messages to receive attention), search, activation (getting ready to deal with the event), and set (preparation to respond in a certain way) (p. 6). Gibson (1970) isolates three attentional processes involved in extracting invariant information from the variable flux: perceptual abstraction of information from the context, filtering irrelevant aspects of stimulation, and active exploratory search. Filtering as a perceptive mechanism for reducing noise is the object of some dispute. First proposed by Broadbent (1958) and subsequently revised by him (1967, 1970) and by Triesman (1964) to account for the fact that information supposedly filtered out may reach the subject's attention under certain circumstances, the concept has been opposed by Deutsch and Deutsch (1963) who maintain that all stimuli reaching the senses are analyzed for meaning. Other writers view selective attention as testing and remembering one set of anticipations over another (Hochberg, 1970), or as the allocation of cognitive resources to a limited segment of the stimulus field (Neisser, 1967). Complementing his account of selective attention, Neisser presents a "fragment theory" to explain veridical perception from incomplete stimulus information; he suggests that set, familiarity, and context predispose the organism to perceive one stimulus configuration over another.

All these variables have implications for reading, and have indeed been discussed in relation to the perceptual mechanisms of the adult reader. Much of the work on perception cited in this section is relatively new and refers to the perceptual processes of adult subjects, but we do not know how well these descriptions fit the perception of the young child. Bower's work suggests that the infant's perception begins to approximate that of the adult in some respects at quite an early age, but we cannot be certain that this is true in all respects. The variables of set, familiarity, and context advanced by Neisser are likely to change with the child's developing cognitive and linguistic competencies. Further studies of both visual and auditory perception as it functions through the preschool and elementary school years are needed.

Piaget takes a different view of the early perceptual processes we have been discussing. For him they are not truly developmental because they do not show sequential changes with age. They are "field effects," or basic organizing forces, part of the infant's initial equipment which have survival value for the individual and persist without appreciable change throughout life. By contrast, "perceptual regulations" begin to emerge around the age of three. During the preoperational period, when decentering of both perception and thought occur, the child becomes increasingly able to reverse figure and ground, to integrate parts and wholes and to scan configurations in systematic and novel ways. Perception is centered on the dominant aspects of the visual field, which tend to be overestimated, while the remaining elements are underestimated. Perceptual strategies such as exploration, reorganization, and schematization can compensate in part for the primary deformations, especially as these activities come more and more under the control of operational thought. From

Piaget's theory of perception, it would follow then, that training in the above-mentioned perceptual activities after age 3 would be more valuable than early attention to reading *per se*. Elkind (1970) found that black second-grade children made more progress as a result of such training than a control group which received equal time in regular reading instruction.

Piaget's theory and research thus provide a counterbalance to the premature conclusion that the young child's perception is similar to the adult's, enabling him to accomplish the same tasks albeit at a more primitive level. Work by Vurpillot (1968), and some Soviet psychologists (Zaporozhets, 1969) support the developmental aspects of perception.

## Cognition

At the other end of the sensation-cognition continuum we find that the most comprehensive framework for investigation has been provided by Piaget's theory. The genetic evolution of intelligence has been studied in the development of the child's concepts of conservation, causality, reality, and morality. Unlike his theory of perception, Piaget's theory of intelligence has not as yet been applied directly to reading, the probably reason being that in general, the visual decoding aspects of reading have received much greater attention than the processes of comprehension. One may speculate that the application of Piagetian cognitive theory to reading may open up a highly fruitful field of inquiry. Even at the practical level it seems reasonable to suppose that an understanding of the child's cognitive development in terms of Piaget's concepts would have implications for the kinds of reading materials suitable for different age groups. It would be premature to suggest what these might be, however, until a more complete rapprochement between cognitive psychology and the psychology of reading occurs.

Furth (1970) has made a broader application of Piagets' theory to reading and to education. The school, he believes, has failed in its primary mission, which is to produce citizens who are adept at solving problems. It goes without saying that our society is in desperate need of people who can solve scientific, technological, and social problems. Traditional education, with its emphasis on information gathering and respect for authority seems to be ill-equipped to fulfil this mission. Its task should be to give children opportunities for solving problems, to show them how to find alternative paths to the same goal, and to provide them with the tools for problem-solving. Reading would be such a tool—one of an entire arsenal. Unfortunately, as Furth sees it, the school has chosen to elevate the tools, especially reading, to the status of a major objective. It has lost sight of the end, and has substituted a means to the goal for the goal itself. Furth maintains that a school cannot gear its resources to the teaching of reading and at the same time expect to do an adequate job of teaching problem solving. This may seem an extreme statement to some teachers, but if one sees reading instruction as inexorably tied in with the lock-step curriculum, it becomes more acceptable. Furth expands on his major thesis to suggest a variety of ways in which the teacher may institute a curriculum which emphasizes problem solving by the children.

Thus, two authors, both working within a Piagetian framework, reach different conclusions about the teaching of reading. Elkind proposes perceptual training for some (perhaps all) children as a precursor to reading instruction in the early grades, while Furth seems to view reading as a skill which might well be acquired over the elementary years as an incidental tool for problem-solving.

Interestingly enough, Rohwer (1971), who specifically rejects the Piagetian notion of critical periods, has come to a similar conclusion based on different premises. In a recent article in the *Harvard Educational Review* entitled "Prime time for education—early childhood or adolescence?" Rohwer cites cognitively oriented preschool programs as the only kind which have produced demonstrable longstanding gains in achievement. However, his conclusion is not, as one might expect, that more programs of this kind are needed. On the contrary he maintains that very little of present-day elementary education is relevant to life outside the school, and should be radically changed to incorporate skills of discrimination, classification, communication, and problem-solving. All these skills, including the 'sacred cow' of reading, should be learned not at a particular age laid down by society, but at the time the child can acquire the skills (and the prerequisite subskills) readily and successfully, a conclusion not too far removed from the general position assumed by Furth.

A consideration of some recent thinking in children's cognitive development thus presents us with a situation in which the reading teacher may derive two different conclusions from the same theory, or a similar conclusion from two different theoretical standpoints.

## LANGUAGE

A preoccupation with the child's perceptual and cognitive development may well lead to our placing primary emphasis on the decoding and word recognition aspects of reading. But the work of Chomsky and others on the generative nature of language has drawn attention to the role of linguistic and information-processing skills in reading. Kolers (1970), in an article entitled "Reading is only incidentally visual," cites evidence from several studies to show that good readers are faithful not to the words they see printed but to the substance of the messages the words convey. Words are not neutral graphic stimuli awaiting translation into associated phonemes, for in order to identify them, one has to know something about them, e.g. that they belong to a certain language. Moreover, there is no necessary serial sequence in the rapid reader's scanning of text. He has mastered the are of selecting clues which enable him to process and assimilate the information directly into his own cognitive structures. In Smith's (1971) terms, the skilled reader goes directly to "immediate meaning", whereas the less fluent reader resorts to the "lower route" of "mediated meaning identification," which interposes a step of word-by-word identification. The implication is that reading instruction should move away somewhat from emphasis on visual recognition, and concentrate on search for clues and information-extracting skills, a suggestion which is congruent with Rohwer's idea of postponing reading until it is relevant to life tasks.

On the other hand, theories of language acquisition which were developed in the 1950s and 60s tended to emphasize the early formation of a complete grammar. Thus as late as 1966, McNeill (1966) was writing: "The fundamental problem to which we address ourselves is the simple fact that language acquisition occurs in a surprisingly short time. Grammatical speech does not begin before 1½ years of age; yet as far as we can tell, the basic process is complete by 3½ years [p. 15]." Lenneberg (1967) also sees the period from 2-4 years as critical for language learning, (although he does see primary language acquisition continuing until adolescence). From this fact one might be tempted to draw either of two conclusions. The first would be that the process of language acquisition is largely irrelevant to reading, since it is virtually complete before reading begins (Wardhaugh, 1971). The second would again emphasize the importance of early reading activity to capitalize on the rapid growth of language.

More recent work, while not discounting the importance of the early period, has restored some emphasis to the continued learning of grammar throughout the elementary school years. Numerous investigations have shown that significant language development occurs after age six. C. Chomsky (1969), for example, points out that several grammatical developments occur after age 6, a most striking example being growth in the use of pronouns. O'Donnell, Griffin, and Norris (1967) found acquisition of new transformations between grades 3 to 5 and 5 to 7, while Menyuk (1964) found examples of more complicated structures as age increased.

Hence inferences about reading based on the earlier position of the linguists, are no longer tenable in the light of this more recent research.

Another aspect of language development which has current salience for reading practice is the issue of dialect. Many linguists appear to be discarding the notion of language deficit which was in vogue about five years ago. Rather they see the inner-city child as having a language system which is well developed both syntactically and semantically, which has some overlaps with standard English, but also many differences. Since most school texts and other reading materials are in standard English, the issue becomes that of finding the most effective ways of introducing the dialect speaker of these standard materials. The various alternatives have been discussed by Wolfram (1970). If extant materials are retained, one may teach standard English, prior to reading, and the child may be asked to render these materials in standard English or in dialect form. The latter requires that the teacher be thoroughly familiar with the dialect in order to distinguish between the dialect rendition and genuine reading errors. If materials are revised, one may eliminate all features which may cause problems for the lower-class speaker (e.g. the possessive 's), or one may construct beginning materials in dialect form with gradual transition to standard text. Wolfram concludes that, in spite of the outspoken rejection of dialect readers by some member of the black community, "the magnitude of the reading problem suggests that experiments must be made, with alternatives which may involve the potential changing of materials and curricula [p. 22]." Another possible approach is to mix standard- and dialect-speaking children in the preschool to permit the latter to gain an understanding of standard English even though they do not use it in speech. While this may be a preferable alternative, it is probably not feasible on a large scale at this time.

On this particular issue, the educator may find that, however well-grounded in linguistic theory his inferences may be, he cannot translate them into practice without taking cognizance of the social and political context in which the school operates.

### MOTIVATION

Unfortunately there is no technical sense of "affection" corresponding to use of the term "cognition": If there were, it would more accurately describe the complex of factors which might be considered under the present rubric.

The study of affective factors and their relationship to cognitive development has undergone considerable change in the last decade. In one sense with the current disillusionment with formal education, motivational aspects have come into prominence in such forms as humanistic education, sensitivity training, alternate universities, etc. On the other hand, the boom in cognitive psychology has deflected much research energy from the study of nonintellectual factors. If one wanted to trace the development of this movement away from the effective toward the cognitive, no doubt White's (1959) classic paper in which he elevated competence to the status of a primary drive would stand as a landmark. But the major source of the change is probably to be found in the current absorption with Piaget's cognitive theory. Piaget does not disregard the motivational aspects of thought; on the contrary he seems to consider motivation as an inherent and inseparable dimension of thought, in the sense that it is part of the ongoing process of intellectual activity. Perhaps the perennial difficulty of measurement in the affective domain is also partly responsible. Whatever the reasons, the fact remains that research on nonintellectual factors in learning seems to have suffered a relative decline. It is interesting to note that the 600-page report of the Literature Search project (Davis, 1971) which addressed itself to every facet of reading made only passing reference to the entire affective domain—a gross omission in my estimation. (The one exception being the paper by Entwisle, referred to below.)

I believe that there is a dimension here which is perhaps more difficult to take hold of, and therefore less rewarding in immediate payoff, but which nonetheless demands continuing attention. I have reviewed the literature on affective factors and reading elsewhere (Athey, 1970). Although it is plentiful, one missess the connecting thread of a good theory to make sense of the plethora of inconclusive and contradictory data. Perhaps such a theory, when we find it, will address itself to the changing motivations and how they affect learning at each age of the lifespan. Meanwhile, it is to be hoped that we will not lose sight of this important domain of inquiry.

Such a theory would need to explain the effects on learning of cultural and social class differences. These have been well-documented in relation to language and reading by Entwisle (1971), and include such variables as control beliefs (knowledge of one's ability to manipulate the environment to meet one's needs), self-confidence, inflexible family role learnings, and the like. Entwisle concludes from her review that "we can only meaningfully teach reading to lower class children when it begins to make sense of their lives [p. VI, 145]."

## CONCLUSION

In this paper, I have discussed some recent literature in perception, cognition, language, and motivation as they pertain to reading. Such a review must of necessity be highly restricted in each area, but I chose to touch on all four areas deliberately, to make the point that the reading teacher, as well as the professor of education, needs not only to be cognizant of what is happening in these (and other) fields, but to keep some perspective among them. If we fail to keep this balance, we are at the mercy of temporary fads which offer panaceas on the basis of limited data.

If education is to be a science, it must progress in the same way other sciences do, by acquiring a systematic body of knowledge, rather than by chasing every fad which claims to have some scientific basis. Education is particularly vulnerable to faddism, because its sources of input are more numerous, but at the same time disparate. Consequently, educators need to be in one sense more open to new developments, and yet more defensive. They must scrutinize each innovation for its scientific underpinnings, and determine whether these conflict with what is known in other fields.

We lost the 'whole child' in the 1950s, and are rebuilding him scientifically from our accumulated wisdom. Putting him together again promises to be a long and arduous, but rewarding process.

## REFERENCES

Athey, I. J. Affective fatcors in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading.* Newark, Del.: International Reading Association, 1970. Pp. 98-119.

Bower, T. G. R. The visual world of infants. *Scientific American*, 1966, *215* (6), 80-92.

Bower, T. G. R. The object in the world of the infant. *Scientific American*, 1971, *225* (4), 30-38.

Broadbent, D. E. *Perception and communication.* London: Pergamon Press, 1958.

Broadbent, D. E. Word frequency effect and response bias. *Psychological Review*, 1967, *74*, 1-15.

Broadbent, D. E. Stimulus set and response set: two kinds of selective attention. In D. I. Mostofsky (Ed.), Attention: contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970.

Chomsky, C. S. *The acquisition of syntax in children from 5 to 10.* Cambridge, Mass.: M. I. T. Press, 1969.

Davis, F. B. *The literature on research in reading with emphasis on models.* New Brunswick, N. J.: Rutgers University, 1971.

Denis-Prinzhorn, M. Perception des distances et constance de grandeur (etude genetique). Archives Psychol Geneve, 1961.

Deutsch, J. A., & Deutsch, D. Attention: some theoretical considerations. *Psychological Review*, 1963, *70*, 80-90.

Elkind, D. Reading, logic, and perception. In D. Elkind (Ed.), *Children and adolescents.* New York: Oxford University Press, 1970.

Entwisle, D. Implications of language socialization for reading models and for learning to read. In F. B. Davis (Ed.), The literature of research in reading with emphasis on models. New Brunswick, N. J.: Rutgers University, 1971. Pp. 6/101-158.

Fantz, R. L. Visual perception and experience in infancy: issues and approaches. In F. A. Young & D. B. Lindsley (Eds.), *Early experience in visual information processing in perceptual and reading disorders.* Washington, D. C.: National Academy of Sciences. 1970. Pp. 351-381.

Furth, H. J. *Piaget for teachers.* Englewood Cliffs: Prentice-Hall, 1970.

Gibson, E. J. *Principals of perceptual learning and development.* New York: Meredith Corp., 1969.

Gibson, E. J. The development of perception as an adaptive process. *American Scientist*, 1970, *58*, 98-107.

Haber, R. N. *Information-processing approaches to visual perception.* New York: Holt, Rinehart and Winston, 1969.

Hochberg, J. Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention: contemporary theory and analysis.* New York: Appleton-Century-Crofts, 1970.

Kolers, P. A. Reading is only incidentally visual. In K. Goodman & J. T. Fleming (Eds.), *Psycholinguistics and the teaching of reading.* Newark, Del.: International Reading Association, 1970. Pp. 8-16.

Lenneberg, E. H. *Biological foundations of language.* New York: Wiley, 1967.

McNeill, D. Developmental psycholinguistics. In F. Smith & G. A. Miller (Eds.), *The genesis of language: a psycholinguistic approach.* Cambridge, Mass.: M. I. T. Press, 1966. Pp. 15-84.

Menyuk, P. Alteration od rules in children's grammar. *Journal of Verbal Learning and Verbal Behavior*, 1964, *3*, 480-488.

Miranda, S. B. Visual abilities and pattern preferences of premature infants and full-term neonates. Unpublished manuscript. Case Western Reserve University, 1969.

Moray, N. *Attention: selective processes in vision and hearing*. New York: Academic Press, 1969.

Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

O'Donnell, R. C., Griffin, W. J., & Norris, R. C. A transformational analysis of written and oral grammatical structure in the languageo f children in grades 3, 5, and 7. *Journal of Educational Research*, 1967, *61*, 35-39.

Olson, W. C. *Child development*. Boston: D. C. Heath, 1949.

Rohwer, W. D., Jr. Prime time for education—early childhood or adolescence? *Harvard Educational Review*, 1971, *42*, 316-341.

Rosenblum, L. A., & Cross, H. A. Performance of neonatal monkeys on the visual cliff situation. *American Journal of Psychology*, 1963, *76*, 318-320.

Smith, F. *Understanding reading*. New York: Holt, Rinehart and Winston, 1971.

Triesman, A. M. Selective attention in man. *British Medical Bulletin*, 1964, *20*, 12-16.

Vurpillot, E. The development of scanning stratagies and their relation to visual differentation. *Journal of Experimental Child Psychology*, 1968, *6*, 632-650.

Walk, R. D. The development of depth perception in animals and human infants. In H. W. Stevenson (Ed.), *Concept of development*. Society for Research in Child Development Monograph No. 107, 1966, *31*, 82-108.

Walk, R. D., & Gibson, E. J. A comparative and analytic study of visual depth perception. *Psychological Monographs*, 1961, *75*(15).

Wardhaugh, R. Theories of language acquisition in relation to beginning reading instruction. *Reading Research Quarterly*, 1971, *7*, 168-194.

White, R. W. Motivation reconsidered: the concept of competence. *Psychological Review*, 1959, *66*, 297-333.

Wolfram, W. Sociolinguistic alternatives in teaching reading to nonstandard speakers. *Reading Research Quarterly*, 1970, *6*, 9-33.

Zaporozhets, A. V. Some of the psychological problems of sensory training in early childhood and the preschool period. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary soviet psychology*. New York: Basic Books, 1969. Pp. 86-120.

JOHN J. GEYER
Rutgers University

# MODELING THE READING PROCESS: PROMISE AND PROBLEMS

A decade ago, at the National Reading Conference, Kingston presented a paper which traced briefly the lack of consensus regarding the nature of reading ability, a lack which he felt was responsible for the little advancement in measuring and teaching reading which had occurred over the previous forty years. Kingston concluded that:

> It seems likely that we must seek bold new conceptual frameworks and theoretical designs if we are to make progress. We now give lip service to the idea that reading is a highly complex process. We have discovered that we cannot adequately define it by merely listing its observable attributes. Simularly we cannot explain it by using equally complex and abstract terms [such as "thinking"]. A possible solution to our dilemma may be found . . . in the use of operational definitions such as those employed in the physical sciences and in current personality research. Such an approach would have the advantage of providing a theoretical framework in which components of reading could be treated separately (1971, p. 103).

Kingston recommended the development of models as a method helpful in describing and clarifying complex phenomena. He believed that modeling ". . . has the advantage of enabling one to reason in operational terms and to formulate certain conditions under which various components function or operate in the reading process [p. 103]."

In 1971, just eleven years later, we find the situation far different.

During the past year, a number of scholars have been reviewing the literature on models of the reading process as part of the USOE's Targeted Research and Development Program in Reading.[1] This work is now complete and detailed reports are available in the final report, entitled *The Literature of Research in Reading with Emphasis on Models*, edited by the Project Director, Frederick B. Davis (1971).

Ninety-seven references were evaluated by the reviewers that were categorized as presenting comprehensive or partial models of the reading process or of processes related to reading. Included within these references were at least seventy-seven distinct works classified as models. Eight of these were classified as comprehensive models of the reading process. The remainder were comprehensive and partial models of processes involved in or related to reading. Only three of the ninety-seven references were published prior to 1960. Of the reading process models only Holmes' Substrata Factor Theory (1955, 1970) predates 1960, the next earliest being my own model first presented in 1966 (1966, 1968).

In 1971, then, we are faced with an embarrassment of riches. Not all of the models are equally useful, of course, and many represent processes whose relationships to reading are quite remote. In a section of the Davis Report, I reviewed forty-eight models which seemed to me to be the most useful (1971a). Even this reduced corpus, however, includes a bewildering array of complex models requiring a sophistocated understanding of a number of disciplines to assay. This corpus comprises a large body of theory, empirically based and recent, rich in both promise and problems.

Some of the major problems associated with this large body of models are the focus of a paper contributed to the Davis Report by Kingston (1971). The issues and questions raised by Kingston must be met if the promise of this new body of knowledge is to prove useful. In the remainder of this paper, I will discuss these problems and suggest the conceptual strategy which seems to me most promising in meeting them.

## MODEL TYPES AND DEFINITIONS

As Kingston states, the word "model" 'is one of today's most popular words (1971, p. VIII-61). It would be difficult to formulate a definition abstract enough to include all things called models. This is a problem of some concern to philosophers of science and considerable efforts on this problem have been devoted to the sorts of things philosophers do. Yet this seems to be the type of problem best left to the philosophers and I would not wish to see a moratorium on modeling until the philosophers reached agreement.

---

[1] Scholars involved in the review of models of the reading process were Frederick B. Davis, Robert Efron, Albert Kingston, Paul Kolers, Jane Mackworth, Norman Mackworth, Karl Pribram, Richard Schiffman, Stanley Wanat, and Wendall Weaver. The writer was coordinator of this area and Associate Director of the Literature Search Project.

I personally like the distinction drawn by Kingston in the 1960 paper. A model is a form of reasoning. It represents a preliminary stage of thinking and it is expected to require alteration and expansion. It does not explain all of the many questions concerning a phenomenon, but it has the virtue of combining many known facts into a broad conceptual framework. Seen in this way, a model is a working approximation of the more formal theory.

To the serious student of a process as complex as reading, some device which allows a tentative conceptualization of the whole is an ever-present necessity if he is to avoid recapitulating the history of experimental psychology by becoming absorbed into the fine points of light-time relationships, visual masking effects, retinal hemifield parameters, fixed retinal images, Maxwellian views and the like. He must keep in mind constantly that his field is understanding *reading*, and that such fine points are means to that end, not ends in themselves. A model is a very useful tool for this purpose, and the relationship this form of modeling might bear to methematical formula, computer programs, and other reasoning methods also called models is not an overriding question. Neither is it a question which will go away, however, for it creates subsidary problems which will be discussed below.

## Level of Rigor

One of the most serious problems facing the scholar who uses modeling as a method of reasoning is the question of the level of confidence at which one publishes. The nature of the modeling process ensures that all elements of a model will not be equally verifiable. One may have considerable evidence and confidence in some components, e.g., short-term memory components, but little more than a feeling of logical necessity about others. One can publish a tight paper on short-term memory in an appropriate journal and this should be done. But such papers will contribute little to our understanding of reading until someone relates them to the larger process, and this is the function of the model.

Consequently, there is a wide difference in level of rigor employed between components of the same model and between models. These differences are not too serious in the basic contributing disciplines where the level of technical sophistocation of the readership and traditions of vigorous controversy are controlling factors.

In an applied field such as reading, with a tradition of presenting all ideas as final scientific truth, the problem is serious indeed. It is remarkably easy to tack together a composite of borrowed components and rank speculations buttressed by this or that reference into a new model. Such models have already been produced and more can be expected momentarily. All the possible answers to this problem which I can think of tend to be hopelessly pious exhortations. Perhaps a modeling fad will at least expose to view the level of theorizing which has been true of reading for some years. When it passes, quickly as do all educational fads, it may leave in its wake a residue of scholars seriously interested in the ideas. In the meantime, let the buyer beware. All PERT charts, wiring diagrams, and computer programs should be judged with extreme caution and all components not accompanied by a full discussion of the evidence should be viewed as someone's guess. Be particularly wary of models which predict only that evidence from which they were constructed. And be suspicious of any model which springs forth full-blown, as if from the forehead of Zeus. Miraculous conception is as rare in this field as in any other.

### DISJUNCTIVE CATEGORIES

The major criticism in the Davis Report paper contributed by Kingston was that the models represent disjunctive categories. He states:

> One can, quite correctly, ask the question: What relationship does the model have to the reading process as commonly seen? The variation among so-called partial reading models attests to the diversity in background and interests of those concerned with reading. Language acquisition and utilization models, psycholinguistic models, information-theory models, perceptual models, learning models represent disjunctive categories. There is an urgent need to bridge these categories or to unite them in a logical manner. Right now it is doubtful if an Aristotle would understand the cumulative effect of these models. Certainly, they fail to help us understand the complexities of the reading process (1971 p. VIII–62).

While I can easily emphathize with the distress this paragraph represents, I find myself in disagreement with the details. I think the body of models has done much to help us understand the complexities of the reading process, if only to remove them from the lip service level deplored by Kingston in 1960. Those complexities are now explicit—they can be reviewed, considered, tested. It can be confidently predicted that the models will have to become much more complex before they bear much relationship to the reading process as commonly seen. But then, we have long known that reading is a complex process and that merely substituting equally abstract terms such as thinking does not explain it.

I would also disagree that model categories are disjunctive, at least in the sense that word is employed in logic, *viz.*, two or more alternatives, only one of which can be true. It seems feasible that models of the reading process could be developed from a number of viewpoints which would not only be valid but which would supplement each other in our understanding of the whole. But perhaps Kingston meant the term disjunctive in its entomological sense, *viz.*, having the head, thorax, and abdomen separated by deep constrictions. This sense of the term comes close to the situation: that the very evident separations between model categories are more structural than functional. In reality all are interconnected in a general system.

The basic problem Kingston calls attention to is, nonetheless, real: seventy-seven models are not a whole lot better than none, if one is looking for immediate application to the classroom. There is a need for the synthesis he calls for, a need which was anticipated from the start. Much synthesis is already reflected in many of the models and this process can be expected to continue as the data base grows.

The task of synthesis, while technical and difficult, will not be as formidable as it appears at first sight. Indeed, I would class the potential for synthesis as one of the major promises of this body of models, rather than one of the problems. A more serious problem for the field will be the need to wait for such synthesis to occur, rather than running off in seventy-seven directions with new reading programs based on some model. I have already been scheduled to take part in a conference program discussing the implications of these models to education.

Many models which are quite similar in their essential components appear more dissimilar than they actually are for two reasons. The first is the frequently accompanying diagram which is drawn to highlight and magnify differences with competing models, and which in different models can make identical components appear unrelated. These diagrams are becoming more standardized, however, as modelers borrow effective techniques from each other. The second, and more fundamental, reason is the use of highly idiosyncratic, even exotic, names for the same basic components. Thus, short term memory components can masquerade under such varied nomenclature as visual image, iconic store, echo box, perceptual image, STM, VSTM, R Buffer, VIS, AIS, IMS, iconic buffer, waiting room, sensory register, logogen, STVM, short-term store, primary memory, short-term reproductive memory, hologram, and many others. While that list includes more than one system, it does represent a very few. This imaginative nomenclature may turn out to be an advantage in the long run, as it prevents synthesis on the basis of clang associations. One must know the research upon which the models are based to understand how many genuine differences these words represent.

Another factor which will ease the task of synthesis is that modelers choose domains of varying size for models of correspondingly varying detail. These differences in levels of models occur within the same category. Thus a model of general information flow with a relatively empty box representing short-term memory might be matched with detailed models of that specific component. This matching must be done on the basis of the research from which the models were generated—one important function of reporting one's reasoning in model form is to guide the reader to this research.

### A CONCEPTUAL STRATEGY FOR SYNTHESIS

One danger in attempting to synthesize a number of models is that each such attempt may only add a new model to the pile. This danger will be least if synthesis proceeds slowly and naturally as a function of normal scholarly activity. This activity should include full credit for borrowed ideas and would be strengthened if the use a borrowed model was put to had the concurrence of the original modeler. At the very least synthesis should be accompanied by considerable comunication between scholars. Since this communication will be between specialists in a variety of disciplines, not the least difficult task will be the development of approaches and terminologies that will allow these scholars to interact meaningfully.

In reviewing the body of models for my contribution to the Davis Report, I was impressed with the potential of information-processing models to cut across the boundaries of traditionally discrete disciplines. All such models, and they comprise the bulk of the models related to the reading process, are concerned with the flow of information through the organism. This central concern creates a parallelism between models, whether written from a psycholinguistic or a neurological point-of-view. This parallelism, in turn, creates such a strong tendency toward synthesis that such models frequently mix evidence from a variety of disciplines in their discussions of components.

In reviewing the information-processing models, I gradually became aware that the basic discipline from which the model was written was a relatively unimportant distinction, but that there was another dimension which could serve fruitfully as a basis for analysis and synthesis. This was the degree to which the model dealt with one or more of three concerns. These concerns are Type S, the programs by which the stimulus is analyzed to yield useful information; Type O, the organismic systems and operations involved in processing information; and Type N, the neural substrates underlying components of behavior. These three types of concerns, while logically and experimentally separable, are so complementary that the distinctions between them have been blurred in most models.

In a paper delivered to the American Psychological Association this fall (1971b), I briefly described the S, O, and N categories. From reactions I have received, it is apparent that two sources of misunderstanding occurred which can be clarified by discussing these categories more fully. First, the categories are not intended, and will not serve usefully, as a taxonomy for the classification of models, since most models mix these categories. Such mixtures are not always successful, the modeler inevitably being stronger in one category than another.

A second difficulty arose due to the similarity of the categories S, O, N with the SOR distinctions of classic psychology. For that reason I shall add a "prime" to the S and O of the categories suggested here. The distinction is fundamental. If we were to represent a SOR situation as a rectangle with the long axis representing time, we could divide the rectangle into its SOR components with two vertical lines. The same rectangle would be divided into its S'O'N components by two *horizontal* lines. In other words, S' models are concerned with all that happens to the information presented in the stimulus throughout the SOR situation. O' models seek to delineate the organismic systems and their operating characteristics which are utilized during the flow of information.

Models representing pure forms are rare, but three such exist for the initial input stages of visual perception. Gibson's work on the analysis of critical features in letter recognition is concerned with the stimulus characteristics extracted by the organism to effect recognition (Gibson, 1970; Gibson). As such, it falls within the Type S' category. Noton and Starks' scan path hypothesis is concerned with the operation of the attention components in vision and is a Type O' model (Noton, 1970; Noton & Stark, 1971). The neurological model of the orienting reflex by Sokolov (1969) is a Type N model. These three models are in no sense disjunctive or competing, but to the extent they are valid, complement each other in presenting an understanding of the total input process.

As a strategy for synthesis, the development and synthesis of information-processing models in the three categories suggested holds much promise for the understanding of reading. Central to this task is the development of the O' model. The ramifications of a comprehensive and valid Type O' model to reading measurement, diagnosis and remediation are rich and almost unexplored. Combined with the Type S' and N models, a truly diagnostic-prescriptive approach to individualized reading instruction could become more than an idealistic slogan. It cannot happen tomorrow, but it will not happen at all unless we get on with the task.

## REFERENCES

Davis, F. B. (Ed.), *The literature of research in reading with emphasis on models.* New Brunswick, N. J.: Rutgers University, 1971.

Geyer, J. J. Perceptual systems in reading: prediction of a temporal eye-voice span constant. Final report OE 5-10-362. Cooperative Research Program, U. S. Office of Education, 1966.

Geyer, J. J. Perceptual systems in reading: the preditcion of a temporal eye-voice span constant. In H. K. Smith (Ed.), *Perception and reading.* Newark, Del.: International Reading Association, 1968. Pp. 44-53.

Geyer, J. J. Comprehensive and partial models related to the reading process. In F. B. Davis (1971, V 1-51). (a)

Geyer, J. J. An eclectic model of the reading process. Paper presented at the meeting of the American Psychological Association, Washington, D. C., 1971. (b)

Gibson, E. J. Perceptual learning and the theory of word perception. Paper presented at the meeting of the American Psychological Association, Miami Beach, 1970.

Gibson, E. J., Osser, H., Schiff, W., & Smith, J. An analysis of critical features of letters tested by a confusion matrix. In H. Levin, E. J. Gibson, A. L. Baldwin, J. J. Gibson, C. F. Hockett, H. N. Ricciuti, & G J. Suci (Eds), *A basic research program on reading*. Ithica: Cornell University, 1963.

Holms, J. A. The substrata-factor theory of reading. In A. J. Figurel (Ed.), *Reading and inquiry*. Newark, Del.: International Reading Association, 1965. Pp. 317-319.

Holmes, J. A. The substrata-factor theory of reading: some experimental evidence. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, Del.: International Reading Association, 1970. Pp. 187-197.

Kingston, A. J. A conceptual model of reading comprehension. In E. P. Bliesmer & A. J. Kingston (Eds.), *Phases of college and other adult reading programs*. National Reading Conference 10th Yearbook. Milwaukee: the Conference, 1961. Pp. 100-107.

Kingston, A. J. Disjunctive categories in ephemeral models. In F. B. Davis (1970, VIII 61-66).

Noton, D. A theory of visual pattern perception. *IEEE Transactions on Systems Science and Cybernetics*. (Vol. SSC-6), 1970.

Noton, D., & Stark, L. Scanpaths in eye movements during pattern perception. *Science*, 1971, *171*, 308-311.

Sokolov, E. N. The modeling properties of the nervous system. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary soviet psychology*. New York: Basic Books, 1969. Pp. 671-704.

PATRICIA ELIAS
Educational Testing Service, Berkeley

# A CRITICAL REVIEW OF THE INFORMATION BASE FOR CURRENT ASSUMPTIONS REGARDING THE STATUS OF INSTRUCTION AND ACHIEVEMENT IN READING IN THE UNITED STATES

The somewhat cumbersome title embraced the work done in the Targeted Research and Development Project No. 3. This was one of three projects funded by the U.S. Office of Education as first steps toward building a coherent foundation upon which realistic goals for the Right to Read effort could be based. The charge of the Office of Education to Project 3 was to determine whether or not an accurate appraisal of the status of reading achievement and reading instruction in the United States could be documented by an examination of existing literature in the field of reading. Project 3 focused on three sub-areas of the literature about reading and addressed itself to three principal tasks: the determination of the extent and distribution of the national reading problem; the determination of the frequency of use and distribution of the instructional methods, approaches, materials and equipment for reading instruction; the description of the nature and extent of current practice in the training of those who teach children to read.

To determine if accurate appraisals could be made concerning these issues, the project staff had to resolve two major problems. One was to develop and implement a comprehensive and systematic method for the selection of the literature for the bibliography of the project. The other was to design and implement a procedure for the systematic review and analysis of that literature. Guidelines for both processes were those provided by the writing of William Gephart in the background report for all three projects. The procedures he proposed for a literature search afforded direction for the selection of materials to be reviewed in this project. Selected sections of those procedures, as well as the elements he proposd for profiling the methodological adequacy of completed research, were used as the foundation of the review process itself.

Two professional committees were used to guide the collection of materials for review and to monitor interpretations of evidence. The first, a "technical review" committee, composed of staff members from Educational Testing Service, designed, implemented, and monitored a documents review process, and developed procedures for the selection of the literature to be reviewed. This committee included research specialists, an experienced public school educator, and an experienced teacher trainer. The second, a "logic committee", composed of acknowledged experts in fields related to the tasks of the project, provided a combination of professional judgments which shaped the literature search and selection process. A third interrelated unit was the group of reader-evaluators who did the actual critical review of all documents.

The preliminary bibliography searches involved a computer sweep through all the ERIC documents to date, a hand search of the last ten years of the Education Index, Psychological Abstracts, Sociological Abstracts and Dissertation Abstracts. Searched as well were the publications of government agencies, state departments of education, selected educational research institutes and individuals, foundations, publishers of both reading materials and tests, teacher training institutions, and selected local school districts.

All references developed through the preliminary selection procedure were submitted to the advisory group of reading experts. Members of this committee evaluated the preliminary set of references according to what they believed about the quality of the data and the probable relevance of the references to one or more of the specific research tasks. These evaluations were then combined and a consensus as to inclusion was reached.

The general policy for accepting a document for review was that no reference was included in the bibliography without having been evaluated and ranked by all members of the Logic Committee, although some references were included on which there was not total agreement as to worth or relevance. In following these procedures we hoped to obtain a relatively unbiased body of documents to form the data base for the project.

Monitoring of the developing bibliography by the staff occurred continually throughout the project and took several forms. Every annual review of reading research published in the last decade was searched for references not already included in the preliminary bibliographic lists submitted to the Logic Committee. The same process was used to examine and select citations from occasional reviews of the literature, documents appearing in the published ERIC bibliographies after the initial computer-assisted search, and other bibliographies provided by educational and research organizations.

A second form of monitoring employed was the comparison, midway through the project, of a list of the journals cited in the project bibliography with that of Summers published in the 1968 Reading Research Quarterly. Summers listed the journals which formed the 1956 to 1966 data base for the ERIC collection. Any journals of 40 included most frequently in the ERIC data base which had not appeared in our project bibliography were searched for relevant articles.

Actual critical review of all documents selected by the Logic Committee was performed by a Committee of Readers. These were doctoral candidates in reading, educational psychology, and sociology, and one law student from the University of California, Berkeley. In their review the readers used a single review format which permitted them to make judgments concerning the adequacy of th data reported, and the degree to which the reported characteristics, conceptualizations, methodology or status could be determined from the evidence presented in the literature. The review format was dveloped by members of the ETS research staff in the Western Office who were responsible for writing the final state-of-the-art reports. This group of authors was also responsible for training the readers in the use of the review format. Information included in the format was both objective and subjective. In addition to reviewing a document for specific population or environmental characteristics which were relatively easy to refine, readers were asked to make judgments concerning data generation quality and representatives, as well as the appropriateness of data analysis according to Gephart's model for review. Reviewers also included relevant data tables and an abstract.

Two hundred documents from the project bibliography were reviewed independently by pairs of readers as a basis for obtaining reliability estimates among the project readers. The variables selected for examination were those on which reader agreement was necessary for either the operational or analytical aspects of the project. The reliability co-efficients ranged from .62 to .89.

Every document reviewed for the project was subjected to the same reviewing process in which it was judged for relevance, assigned and specifically related to the tasks and subtasks of the project, profiled as to the adequacy of the research, and summarized. Of the thousands of documents considered, 1,855 were selected to be reviewed for the project.

It is my privilege to present papers authored by Abraham Carp and Richard Harsh. Dr. Maxwell [the next paper] will present the paper she co-authored with George Temp. It is my understanding that ERIC/CRIER, possibly in conjunction with the International Reading Association, will soon publish sections of the report. The complete report (Corder, 1971) will be available on microfiche in the ERIC collection.

Carp in his chapter, *The Reading Problem in the United States,* reviewed the literature published in the United States between 1960 and 1970 to determine the extent of the reading problem in the country. However, the charge of the USOE to identify, analyze, and summarize existing survey and test data which indicate the reading aiblity of various populations in relation to the *individual and social needs* of the populace presented some problems.

Prior to the review, it was assumed that a body of literature existed which documented the reading ability of various subpopulations as a function of sex, age, race, socioeconomic status, etc., and that the various standards of reporting utilized in this literature could be equated to individual and social needs. In the actual review process this assumption was *not* met in fact. There exists neither a good estimate of the reading ability necessary to function satisfactorily in modern society nor a satisfactory estimate of the absolute reading achievemetn of reasonably defined subgroups in the United States.

United States census data provided much of the information on the educational achievement and state of literacy of various subgroups of the total population. Both the *number* of years of education completed and the *status* of literacy are categorized by a large number of demographic variables. The U.S. census definition of *literacy,* based on self-report, assumes that all persons who have completed five years or more of school are literate. *Illiteracy* is defined as inability to both read and write English, or any other language. According to census data, the *illiteracy* rate in the United States is currently one percent of the population 14 years old and older. In the decade since 1959, the number of illiterate persons has *decreased* despite an *increase* in the number of persons 14 years old and over.

Illiteracy, as defined by the Bureau of the Census, is primarily a problem of years of schooling completed. Illiteracy rates drop as even a limited number of years of school increase. With the almost universal availability of 12 years of free public education in the U.S. today, the "production of new illiterates" is probably at a minimum. *Illiteracy* as a significant problem in the United States today is related to age (77% are over 45), geographic region, and to specific ethnic populations.

It must be cautioned, however, that these statistics are based on self-report, or answers to questions *about* literacy, not to demonstraions of performance. The experience of the armed forces in World War II and later indicated that the standard definitions of literacy did not serve to delineate the limited ability of many servicemen to communicate with their fellow men. Not only those defined as illiterate by reason of limited number of school years, but some servicemen with education beyond this limited point, were found to be incapable of performing military functions involving reading and writing.

There is a serious functional literacy problem in the U.S. when functional literacy is equated to number of years of school completed. The size of the problem depends on the level of educational attainment accepted as the standard of functional literacy. If completion of five years of formal schooling is accepted as the standard, there are some eight million functional illiterates; if eight years of schooling is the standard, there are almost 19 million functional illiterates. And if we accept the completion of 12 years of schooling, there are over 70 million functional illiterates. Whatever criterion we choose, the education deficit is related to age, race and ethnic origin, urban, suburban, and rural residence, and region of residence in the country. For example, of the six million persons over 14 years old who have had less than five years of schooling, about 320,000 are under 25 and about four million are over 55. In terms of educational deficiency in the adult populations, those of Puerto Rican or Mexican origin show the greatest deficit, followed by the Black population.

Another body of literature reviewed for the project included test data, and survey and technical material, in which populations were described in grade equivalent scores based on nationally normed tests of reading achievement. Despite the variations based on different tests with differing content, the level of deficit observed was relatively constant. There is no generally accepted test which can serve as the criterion and no extant calibration of tests to a common frame of reference. But some aspects of the reading problem based on norm-referenced tests permit an estimate of the reading deficit for students who are still in school.

If we accept a grade equivalent of 5.0 as the standard for meeting social and literacy needs, about one percent of those with 12 years of education, 3 percent of those with 10 years of education, 13 percent of those with 8 years of education, and 30 percent of those with 6 years of education will read below this standard.

If we accept the figure of 8.0 as the standard, then 13 percent of those with 12 years of education, 24 percent of those with 10 years, and 50 percent of those with 8 years will fail to meet the standard.

The over-all conclusion derived from this review was that the data base does not exist to permit adequate estimates of the reading problem in the United States in terms of a standard of meeting "individual and social needs." The statement of the problem implies the existence of a standard and first priority should be given to the development of such a standard. The Office of Education has recognized this problem and through targeted Research & Development Project #1, has funded research to create such an instrument and to collect performance data indicating the extent to which various population subgroups meet the standard.

The incidence of reading problems from kindergarten through grade 12 has been demonstrated, but the extent of the problem depends on definitions, measures, and populations described. If the average score is defined as the standard, obviously, half of the population will fail to achieve the standard. Larger proportions of individuals whose families are characterized by financial, housing, education, and other disadvantages will fall below the average. However, until some criterion other than a norm-referenced one is accepted, somebody is going to have to read below average. The way various groups perform in relation to the norm may be changed by various intervention activities, and their relative positions may be changed, but fifty percent will still be below average and approximately fifteen percent of those with 12 years of school will read below the ninth grade level. The current information on the relative deficit of various groups does present, however, targets for intervention and suggests the ones that present the greatest need.

The particular goals, however, are not so easily defined. A current target expressed by some of insuring "a year's growth for every year for every pupil" is probably not attainable. A more modest goal related to performance or criterion-based measures may be achievable. With these types of measures, groups may still be differentiated on the basis of numbers or percents able to reach the standard. They may also be differentiated on the basis of other factors, such as time, or effort, or cost to meet the standard. More effort should be directed to developing performance-based standards for reading in the school population similar to those under development for adults, particularly for the upper school grades. Attempts to develop specific objectives related to reading and to translate these into performance-type measures would provide approprite standards for school-age reading programs which norm-referenced tests do not supply.

Further efforts need to be directed to the economic consequences of reading, particularly in the adult population. While there is ample evidence that measures of academic achievement, including reading, correlate with indices of economic status, it is not necessarily true that improvement of reading will produce economic benefits. Much more needs to be known about the reading requirement of jobs. The requirements that many employers have for school certificates or for reading achievement may have no real relationship to the requirements of jobs. Recently the courts and the Equal Employment Opportunity Commission have required employers to demonstrate a relationship between job performance and educational requirements, achievement, or intelligence test scores in order for these requirements to be valid. Despite the apparent success of various federally and state funded programs to provide instruction in reading to unemployed adults, it has not been demonstrated that improvement in reading is the most critical factor in increased employability. Nor is the information available to determine the cost/benefit ratio of improving reading as compared to the same amount of money spent in some other activity, for example, the cost of changing the reading requirements of forms which people must fill out to "survive" so the forms are more understandable and less subject to error.

The "Right to Read" is a right that every individual should have, but effective reading may not solve all our economic and social problems.

Harsh in his chapter, the *Nature and Extent of Current Pracitces in Educating Those Who Teach Reading*, addressed himself to the determination of whether the explicit characteristics of current practices in preparing those who teach reading could be identified from the relevant literature, and, whether these characteristics could be related to the reading ability of the student population. The elements of a chain were examined beginning with state legislation related to reading instruction, and extending through state certification boards' translation of legislative requirement into teacher education requirements, teacher education programs for graduation and certification, and local practices operating to meet state and district requirements for reading instruction.

To provide a statutory background for the certification requirements of the states, the education code of each state was reviewed to determine the nature and extent of the laws which might pertain to the teaching of reading. Despite what you may think, most states (28) do *not* have laws that mention reading as a specific requirement, although 16 of these states have status which require instruction in specific subjects. Twenty-three states have legislation requiring that reading be taught in the schools. None of the state education codes prescribe specific materials, time, or methods of instruction for reading *or* any other subject. The education codes tend to be quite consistent in assigning the determination of such requirements to the state board of education and/or the governing board of the local school district.

All 50 states and the District of Columbia have statutes which require licensing or certification of teachers. However, almost all states have made provisions to grant probationary credentials for those who have not met the regular credential requirements in the event no fully credentialled teachers are available. Although only one state has a statute specifying the certification of reading teachers, 22 have made provisions for the state board of education to issue "special" certificates for certain subjects or fields as deemed necessary.

The state education codes provide a broad framework for the establishment, government, and operation of the public schools, but generally do *not* make specific requirements for the content, materials, methods, or teacher certification related to the teaching of reading.

*What are the state certification requirements for those who teach reading?* All states require professional school personnel who teach in the public schools to hold certificates issued by the legal authority. In all but four states a minimum of a Bachelor's degree is required for elementary and secondary credentials. While the predominant practice is to "endorse" the fields of preparation on the secondary teaching certificates, such endorsement is not found in the issuance of elementary credentials. The credential requirements are commonly stated in terms of degrees or amount of academic preparation with extensive use of "approved programs" of training institutions as the means by which teachers may be certified.

The literature contains conflicting and ambiguous information concerning the incidence of reading and reading-related courses required for elementary and secondary certification. The extensive use of the "approved program" approach to certification makes it impossible to form definite conclusions regarding the presence or absence of reading requirements.

The review of the existing literature indicates that it is *not* possible to adequately describe the nature and extent of the preparation of elementary and secondary teachers to teach reading from the reports of state certification requirements.

The available literature suggests that states designate specialists in reading by a variety of titles, endorsements, or credentials. The common certification requirement is in terms of courses of particular title or content, with the majority of states also specifying teaching experience. The requirements for all states are described in such a variable and inconsistent manner that explicit conclusions are impossible. The only conclusions that may be made are that no apparent regional trends exist and no common criteria are used by all states for the certification of reading specialists.

*What are the current practices of institutions preparing those who teach reading?* Since the review of the literature revealed no current national summarization of the offerings and requirements of institutions providing teacher education programs, the published catalogs from 374 of these institutions were used as the source of the most current information. Institutional requirements and offerings for teacher preparation in reading were specified in the catalogs of 324 institutions in 49 states and the District of Columbia.

Sixty-four percent of the institutions required a separate reading methods course involving from two to three semester hours of study in the undergraduate preparation of elementary teachers. One-third of the institutions required an integrated reading-language arts course or a general-methods course which included reading for elementary teachers. Twelve percent of the institutions required some form of practical experience or practice teaching concurrent with the required reading courses for the elementary teacher. Only three percent of the institutions listed no reading-methods course requirment for the preparation of the elementary teacher. In addition, 47% of the institutions required those preparing for elementary teaching to take a course in children's literature. All institutions preparing elementary teachers require practice teaching which involved 6 to 16 semester hours generally to be accomplished in the senior year.

The catalog-specified requirements of secondary teachers for preparation in reading were significantly less than the requirements for the elementary teachers. Atlhough only six percent required a reading-methods course, nearly 60% of the institutions offered one or more courses in reading methods at the secondary level.

During the past decade, there appears to have been a very slight change in the requirements of institutions that prepare teachers for elementary and secondary certification. In 1960, as in 1970, the most frequent requirement for certification as a regular elementary or secondary teacher was one course in reading and/or language arts. The institutional requirements for preparation in reading increased slightly for the elementary teacher, while there was a slight decline in the reading course requirements for secondary teachers.

The survey data are not adequate to answer the question of the contents and methods used by institutions to prepare teachers for the instruction of reading. The available information is restricted to courses and semester hours of instruction, and no comprehensive information is available concerning the specific contents, methods and outcomes of the teacher education programs.

*What preparation for teachers or specialists in reading is required or provided by local educational agencise?* A survey of the 20 largest cities in the United States revealed that all of the 17 responding cities employ some type of special reading teacher. There is wide variation in the preparation requirements for those special teachers of reading which commonly do not relate to the state certification requirements. More than half of the cities do not require any academic preparation beyond the regular state certification for the special reading teachers. At the same time, approximately one-third of the cities require or provide inservice education, and 10% of the major cities surveyed have indicated adoption of the IRA requirements for reading specialists. Successful experience as a teacher of reading is a common requirement for those who are designated as specialists or teachers or reading in local educational systems.

Neither the reviewed literature nor the documents provided from a sample of the major cities supplied sufficient information to determine adequately the requirements of local educational agencies for the preparation of teachers or specialists in reading.

The surveys reviewed in the literature suggest that professional school personnel perceive the need for more assistance in providing for students with reading disabilities. The surveys suggested that between 50% and 60% of the students identified as having reading problems were not receiving special assistance. The nature and extent of the preparation for the teaching of reading are far less than the professional reading association recommends *and* less than surveyed teachers felt they needed. The limited surveys of special ethnic and cultural studnt populations provide incomplete but suggestive indication that special problems in reading and language development for these groups are to a great extent receiving little attention. Unfortunately, the literature does not provide comprehensive information concerning the preparation and characteristics of those who teach and provide special assistance in reading to pupils with various needs in all regions of the nation.

The review of the literature failed to provide adequate evidence to answer the important questions of *the relationship between teacher preparation, teacher performance, and student achievement in reading.* The lack of evidence from the literature is not due to a dearth of surveys, writings, and investigations. Rather, it is due to limited or inadequate dseign of investigations which do not account for the multiple variables interacting in a system of education.

Although no definitive conclusions concerning the relationships of teacher preparation, teacher performance, and student achievement can be made, some of the studies suggest that characteristics called verbal facility, flexibility, and empathy for individual learners are more important teacher characteristics than the degrees and courses acquired in academic preparation. Other studies amplify these suggestions with observations that pupils from diverse ethnic and cultural backgrounds may be understood and communicated with effectively in a learning situation by those who understand and share such experiences by virtue of their own background.

The research and analysis reported in the USOE publication, "Do Teachers Make a Difference?" suggest the new developments needed in research concerning educational systems. There appears to be a growing acceptance that future educational research must be designed to explain the model of the system of education that includes input, process, and output variables.

The inability of existing research to provide definitive answers to the major questions addressed in this chapter appears to be partially explained by the design of the investigations which commonly used correlational analyses of single variable of student or teacher characteristics. The principle of parsimony appears to be reflected in the designs of the majority of investigations which have sought the *one* best set of requirements for teachers, the *one* best teaching behavior, or the *one* best method of preparing teachers. Surveys and research studies have commonly identified an array of demographic differences among teachers and students but have not investigated the interactions of the multiple variables in an educational system. From this review of the literature, it appears clear that reliable description and understanding of the complex system of education will depend on future research which investigates the context, input, process, and output interactions among the variables of students, teachers, schools, and society.

Project 3 was one of three inter-related projects funded in 1970-71 as part of the USOE Targeted Research and Development Program in Reading. Several basic assumptions underlie all three projects. It is appropriate to summarize the findings of the papers I have reviewed which relate to two of those assumptions. One assumption was that *improvement in reading seems to have reached a plateau.*

We do *not* think the literature surveyed for this project supports this as a general assumption. The answer depends, in part, on what subquestions are asked to define the main question.

If the question is examined from the standpoint of improvement in reading *achievement,* the work of Gates in 1961 and Schrader in 1968 sugegst, at least tentatively, that the reading ability of pupils in the public schools may have improved over the last few decades. Recent Population Reports of the Bureau of the Census indicate that illiteracy, on the whole, has declined in the United States during several decades, despite a significant increase in the population itself. However, the results of this literature search abundantly support the position that low achievement in reading, and even illiteracy, is a significant problem in segments of the population.

The incidence of reading problems in the United States among both the school age and adult populations (as described in Carp's paper) is related to ethnic and racial group membership, socioeconomic factors, and location of residence. These and other factors, such as age, begin to describe the groups in our population who may be on a plateau of achievement and, thus, appropriate targets for intervention. The generalization that overall improvement in reading achievement has reached a plateau could be supported only when and if optimum effort has been made to effectuate appropriate reading programs for mmbers of these subgroups.

If the problem is examined from the standpoint of improvement in teacher *education,* it is apparent from our survey of the descriptions of courses in reading instruction that there has been little institutional response to the 1961 challenge of Austin and Morrison to improve teacher education. Another assumption was that *summaries of research on reading indicate that most of the research in the field has been done in a manner that prohibits synthesis.*

Those who participated int his project sympathize with the problems of previous synthesizers in the field of reading. We found that the literature typically contains testimonials about programs, or techniques, *or,* reports the results of research in a manner that is difficult to synthesize. Most of the research on method, for example, has been based on the assumption that method alone makes a major difference in learning to read and has ignored, or left uncontrolled, other significant variables such as learner and teacher characteristics. The reader-evaluators who worked in this project and the authors of the papers of this report pointed out that the application of sophisticated statisical analyses to variables that are vague, lack operational definition, and do not reflect the complexity of the reading act, is futile. Our experience tends to confirm the basic assumption that research, *as reported,* has been done in a manner that, if not prohibiting synthesis, does make synthesis difficult and tenuous.

The question posed to Project No. 3 by the United States Office of Education was whether or not it was possible to document from existing literature the status of the several tasks posed to the project. This was to be accomplished by the review of that literature in a manner which uilized a defined and consistently applied set of standards. The standards suggested to be applied were those defined by Gephart as dimensions required of *quality* research. We have applied those standards in an extensive review of the relevant literature. We feel the results will challenge the research community in reading. We hope they contribute to a better understanding of some of the problems involved.

## REFERENCES

Corder, R. *The Information Base for Reading—A Critical Review of the Information Base for Current Assumptions Regarding the Status of Instruction and Achievement in Reading in the United States.* Final Report Project #0-9031, Grant #OEC-0-70-4722 (508).

MARTHA J. MAXWELL
University of California at Berkeley

# RESULTS OF THE SURVEY OF THE LITERATURE ON METHODS AND MATERIALS IN READING

One of the three objectives of Project 3 of the Targeted Right to Read Grants was to answer the following questions concerning the use frequency and use distribution of instructional methods, approaches, procedures, materials and equipment for reading instruction:

a. What methods, materials, approaches, equipment and procedures are used to teach reading in the U.S. and to what extent?
b. What methods of reading instruction are built on essentially different pools of basic knowledge?
c. What relationships between methods of reading instruction and reading achievement of the various subgroups in the population can be shown?

In addition, the research evidence on methods and materials used in teaching reading was critically evaluated.

## PROCEDURE

In order to accomplish the objectives of this study, it was necessary to locate all existing literature that might be relevant and develop a procedure for selecting, systematically reviewing and analyzing that literature. A committee of experts from universities representing diffrent fields in reading and related disciplines screened and selected specific studies, surveys, and other documents from the vast literature in the field and advised on all phases of the project. Each item selected was carefully read and rated by one or more readers. (A completed description of the procedures and rating methods used can be found in Chapter 1 of the complete project report, Corder, 1971.) In addition to rating the item on the specific charges of the project, readers also rated its research quality based on Gephart's (1970) criteria—the representativeness of the sample studied, the adequacy of the description of the treatment used, the validity and reliability of the measurements and the appropriateness of the data analysis procedures. Of the tens of thousands of items screened, 948 were selected by the committee as meriting evaluation for the section on materials and methods. Of these, a total of 244 separate articles were judged by the readers as acceptable (i.e., of high or middle quality.) Eighty articles dealing with methods and 33 on materials were ranked "high" by the evaluators.

## RESULTS

Most of the research studies on reading methods involved the beginning reading stages. 161 of the articles surviving the Gephart criteria were on Grades 1-3, 38 on grades 4-6, 25 on junior high, 2 on senior high, 10 on young adults (college) and 5 on adults. Furthermore, with the exception of "programmed instruction" and individualized instruction," it is rare that the term, "method of teaching reading" is used above grade 3.

The most typical finding reported in the studies reviewed was "no statistically significant differences were found between experimental and control groups." This statement appeared with monotonous regularity. In the studies which reported significant gains at the end of the experimental period, these inevitably disappeared in follow-up studies, and, in some cases, controls exceeded experimental subjcts in subsequent learning. To be sure there were significant differences in some variables in a few studies. At the present time, we lack even a "web of partial evidence" from which teachers can design effective methods for teaching reading.

What the project did reveal were the kinds of problems researchers in reading methods face and the assumptions underlying their efforts.

### *Definition of Reading Methods*

Investigators tended to label rather than to give operational definitions of reading methods. The classification system developed by Chall (1967) was used in grouping methods studies:

1 – Meaning emphasis
2 – Code emphasis
  a. synthetic
  b. analytic
3 – Linguistics
4 – Modified alphabet

5 – Responsive environment
6 – Programmed learning
7 – Individualized instruction
8 – Language experience
9 – Eclectic (author's own or other)

In evaluating the articles, books and other source materials, it was usually possible to fit the author's label of method into one of the above categories, although some researchers only described the materials used in the study. Few studies explicitly described teaching methods and almost none gave information concerning the sequencing of the exact elements in the instructional plan, the type and amount of practice trials for each aspect of the presentation, the kinds of reinforcement used by teachers, nor the total time devoted to the procedure under study. Most of the research merely labeled the "method" used without operationally defining it. In many cases, quite different practices were labeled with the same term (e.g., "individualized instruction" was used to describe programs where students freely chose their own books to read as well as programs where students followed a rigidly prescribed programmed learning instruction sequence; "responsive environment" referred both to the "Talking Typewriter" and to sundry variations on the British Infant School model.) The term, "method" also was applied to quite different phenomena: e.g., classroom organization, overall nature of the instruction, materials used and even to the basic "approach" or philosophy involved. In almost all of the methods studies, materials are inextricably linked with method. For example, in the i.t.a. studies, it is impossible to determine whether the differences reported were due to learning to read in the new alphabet or to differnces in the materials or classroom strategies since students in i.t.a. classes typically used the *Early to Read Series* and controls used, "basal readers." (Warburton and Southgate, 1969)

Studies using "linguistic methods" showed even greater variability in terms of materials used, often appearing indistinguishable from "phonics-based" programs. Inevitably, control groups were taught with the "traditional basal reader" approach. In most studies, the specific basal reader used was not identified nor was descriptive information given as to how the teacher used or supplemented the basal reader materials.

Not only did labels used to describe reading methods have quite different meanings depending on the author, but more sophisticated attempts to classify different reading approaches compounded the confusion. For example, the Bliesmer and Yarborough study (1965) which compared 10 reading methods has been widely cited as demonstrating the superiority of the synthetic approach. A careful inspection of the data they cite reveals the following:

1. The "individualized approach" which they classify as analytic differs significantly from each of the other analytic approaches on at least one comparison.
2. The "individualized approach" classified as analytic did *not* differ significantly on any comparison from the "programmed learning approach" which they classified as synthetic.
3. On 17 out of 50 comparisons, the synthetic methods differed significantly from each other. (See Temp, pp. 128-132, in Corder, 1971, for complete emphasis.)

Chall in her book *Learning to Read: The Great Debate* takes great pains to illustrate that selections from phonics-based readers do not differ from meaning-emphasis materials on meaningfulness, but then proceeds to classify them in different categories. In her analysis of the research she undertakes a deep and scientifically rigorous look at the existing research, but finds herself faced with the problem shared by all synthesizers—i.e., if research quality standards are too high there is nothing left to summarize. Chall resolves the dilemma by accepting studies with recognized major defects on the ground that they are "as good as most studies of their kind." Though she claims that her conclusions stem from her theoretical analysis of the probable course of reading skill, not from "head-counting results," her study still leaves one with the impression that one can merely examine the research literature no matter how inadequate it is and conclusions will emerge.

Authors of textbooks, professors of reading, and investigators who engage in research rituals assume that reading methods exist and that they differ. We have found no evidence to support either assumption.

What the teacher does with materials and lesson plans when he or she closes the classroom door has rarely been reported in the reading research literature. That teachers ignore the controversy on reading methods is suggested from survey data. For example, in the New England Educational Assessment Project (1969), 94% of the first grade teachers reported using basal reader materials. but 94% reported they also used intensive phonics materials. The majority of the teachers said they also used supplementary basal materials, teacher-made exercises, tape-recorders, etc., to teach reading. Ninety percent of the first grade teachers put some or much emphasis on configuration clues (whole word approach), 98% stressed phonics (letter sounds) and the majority of the group also claimed they stressed linguistics (word patterns) and structural analysis. It would appear that teachers use a wide variety of materials and approaches in teaching reading.

### *Other Assumptions Underlying Research in Beginning Reading:*

Implicit in much of the research in beginning reading is the assumption that if a child is started out right at the beginning stages, the advantages he gains over other students in reading achievement will persist permanently. Follow-up studies typically involve periodically administering reading tests in subsequent years without considering the intervening learning experiences or the impact of other factors on reading development. Longitudinal studies tracing the child's reading development and educational experiences are rare indeed. If this research direction continues, one can confidently predict that soon studies will be reported where the achievement of students taught beginning reading in i.t.a. is compared with controls on College Board Examination Scores administered twelve years later.

Another research assumption is that it is good to get children reading early and the faster they learn, the better. This value judgment has no empirical basis and runs counter to our knowledge of the differential patterns of child development and motivational stages in cognitive, emotional and physical areas.

Implicit also in the development of "methods" is the view that if the initial stages of learning to read can be made easier by changing the code, controlling the sequence of presentation of letters of the alphabet, teaching phonics rules or controlling the vocabulary, etc., the child will learn to read more easily and quickly. With the sole exception of i.t.a. research, none of the other investigators have mentioned the problem of transfer, much less studied it. That these attempts may oversimplify the process of reading to a degree that children may not be adequately prepared to make the transition to "real reading involving as it does irregular and polysyllabic words, inconsistent spelling patterns and pronunciation, and complex syntax. It is quite possible that simplifying the beginning reading stimuli may indeed result in negative transfer making it more difficult for some chldren to make the transition to normal written material. Systematic research on transfer is sorely needed.

Underlying most of the research and written materials on methods, despite protestation to the contrary, is the undisguised hope that if only we are careful and scientific enough we will eventually discover the one method of teaching reading that will work with all children.

### *Materials*

There are well over 10,000 separate products to teach reading on the market today including books, filmstrips, multi-level kits, games, etc. This information explosion poses a dilemma for the reading educator who is responsible for making decisions about which materials to use. P. Kenneth Komoski, President of the Educational Products Information Exchange Institute, a consumer's union for school systems and educators estimates that less than ten percent of the educational products on the market today have been field tested, and far fewer have been subject to more rigorous learner verification testing

(Komoski, 1971). Current products for teaching reading being developed by the various regional educational laboratories are being rigorously tested and revised. For example, the First Year Communications Skills Program developed by the Southwest Regional Laboratory (SWRL) has been learner-verified by testing and used with more than 30,000 children in twelve states over a four-year period. The Wisconsin Design for Reading Skills Development (Otto and Askov, 1970) represents another approach in which research has been used to develop interrelated components of basic skills for grades one through six with the goal of aiding schools to make more efficient use of available published materials rather than developing new materials. In evaluating these as well as other systems approaches to the teaching of reading, it is important to recognize that they are designed to *supplement* not *supplant* the regular school reading program.

A comprehensive list of materials "purported to be distinctly phonics programs' 'was compiled by the Educational Products Information Exchange and published as "Product Information Supplement No. 7: Text-based Phonics Programs," in April, 1969. Over 40% of the 189 titles listed were reported to have undergone some sort of field testing.

Few materials designed for high school and college level students have been field tested. An exception is the McGraw-Hill Basic Skills System which includes programmed books for learning a number of reading and study skills and is designed for high school and college students.

Komoski (1971) fears that the proliferation of educational materials has lead to providing schools and teachers with an increasing number of trivial options rather than the high quality alternatives they need. The plethora of educational materials is symptomatic of our technological age which creates a demand for options in all purchases as well as a ned for immediate acquisition of products and information. Teachers, like other consumers, have been conditioned to expect new materials and methods every year and to look to educational technology to help solve classroom problems. (An illustration of this attitude was the incident at the 1970 National Reading Conference when a teacher rose to protest the government's spending money on long-term research when what she needed was something to use in the classroom "Monday morning.")

Field-testing new materials is an expensive proposition. Publishers are unlikely to underwrite this additional production cost voluntarily, particularly since educators do not demand evidence that the product they are purchasing has been tested and, in fact, generally ignore field testing data when they are available.

Even were all materials sold to schools required to have demonstrated learner effectiveness on the nation-wide sample, this would not guarantee that the materials would work equally well with a particular sub-group population in a local school district. School systems need help in establishing criteria for their local needs and in developing inexpensive pilot-testing programs so that materials can be tested on local populations.

### *Use of Materials in Teaching Reading*

We found no recently published, large-scale national surveys on the use and distribution of reading materials in the schools. The local and regional surveys and limited national surveys that we reviewed uncovered no evidence that would refute the findings of earlier surveys that Basal readers are used in almost all elementary classrooms in grades K-6. (Austin and Morrison, 1963, Barton and Wilder, 1964.) However, as mentioned above, regional surveys suggest that most elementary teachers are supplementing basal readers with other materials such as intensive phonics workbooks, teacher built materials, and audio-visual aids—especially tape recorders. Despite the criticism of basal readers expressed by reading experts and the public press, there was no evidence that classroom teachers are rebelling against them nor using them less frequently.

Regional surveys of high school reading programs suggest that machines, multi-level kits, and workbooks are still the most frequently used materials.

The language-experience approach, i.t.a., and individualized instruction, although described frequently in educational publications, actually involve only a very small percentage of classrooms as revealed by the surveys. Computer assisted instruction in reading, perhaps because of its expense, has made few inroads in public schools and is generally limited to a few research projects.

### *The Knowledge Base of Methods and Materials in Reading*

Almost all of the so called methods of teaching reading reported in current literature, including i.t.a., have their roots or counterparts in practices described in the professiona lliterature of the nineteenth century. Proponents of current methods and approaches have selectively draw evidence from the research literature in psychology and linguistics to support their positions. Usually they focus their rationalizations on one particular branch of psychology and ignore or underplay findings from other areas. For example, the rationale for individualized instruction, responsive environment, language experience approaches have in comon an assumption that children can learn to read in a less structured situation than the normal classroom provides. They cite psychological evidence on motivation, differential growth and development patterns and studies on individual differences to support their positions. Analytic and synthetic method proponents cite studies from visual perception and learning to support their positions, but ignore findings which do not support their theory.

*Inter-relationships Between Methods and Materials, Characteristics of the Learner and Reading Achievement*

The research literature dealing with the relationship between characteristics of various sub-populations and materials and methods used in teaching reading revealed little information. Most studies failed to specify the obvious descripters of the population studied—sex, intelligence, ethnicity, or socio-economic background. Only two studies on American Indians survived the Gephart screening system and only four dealing with Asian-American students. In one of the "Indian" studies, the Indian population was comprised of two Indian children.

Most of the studies which specified ethnicity concerned Black, low SES students in ghetto schools. No studies meeting our research criteria were found on middle-class Black children nor on high SES White children with reading difficulties.

If further research is conducted on questions involving school procedures and materials for special groups, researchers should be required to provide adequate data on their populations including sex, ethnicity, ability, grade level, SES background as well as other relevant reading characteristics including physical, emotional and perceptual difficulties.

It may appear to many that the review of over 900 articles on reading methods and materials selected as pertinent to the interests and concerns of this project should yield more information. That it did not was a constant concern to those who put extensive, time-consuming and detailed effort into completing the project. The implications of this study are crystal clear. It is time to refocus the research effort in reading, to raise the standards of designing, conducting and reporting of reading research so that these efforts are compatible with standards of research investigation and reporting that scientists have accepted in developing and communicating their findings. Incomplete, small and partial studies where only the experimenter using his clinical insights can interpret the results may be useful in shaping opnion and developing hypotheses for future research. However, one cannot trust, much less apply information developed from studies that are incomplete, fail to control for significant variables, are based on small populations, do not adequately describe characteristics of the sample, do not specify the conditions under which the study was conducted, fail to specify operational definitions of the variables used (whether they be teacher characteristics, pupil characteristics, methods or materials) and state conclusions that are not supported by the data.

If reading research is to improve, then it is the responsibility of editors of professional journals and other publications, individuals responsible for the awarding of grants and the research consumers to demand more rigorous standards. Researchers must recognize that sophisticated statistical treatments, and elaborate experimental designs will not yield answers to poorly stated questions nor clarify fuzzy concepts.

## REFERENCES

Austin, C., & Morrison, C. *The first R: The Harvard report on reading in elementary schools*. New York: Macmillan, 1963.

Barton, A. H., & Wilder, D. E. Research and practice in the teaching of reading. In M. B. Miles (Ed), *Innovation in education*. New York: Teachers College, Columbia University, Bureau of Publications. 1964. Pp. 361-398.

Bliesmer, E. P., & Yarborough, B. H. A comparison of ten different beginning reading programs in first grade. *Phi Delta Kappan*. 1965, *46*, 500-504.

Chall, J. S. *Learning to read: The great debate*. New York: McGraw-Hill, 1967.

Corder, R. *The Information Base for Reading—A Critical Review of the Information Base for Current Assumptions Regarding the Status of Instruction and Achievement in Reading in the United tSates*. Final Report Project No. 0-9031 Grant No. OEC-0-70-4792 (508). 1971 U.S. Dept. of HEW.

Gephart, W. J. Application of the convergence technique to basic studies of the reading process. Final report. Project No. 8-0737. Grant No. OEC-0-8-080737-4335. United States Office of Education, National Center for Educaitonal Research and Development. 1970.

Komoski, P. K. Statement before Select Sub-Committee on Education, Committee on Education and Labor, United States House of Representatives, May 11, 1971.

New England Educational Assessment Project. Readingi nstruction in New England's public schools. Feb. 1969. ERIC: ED 032 996.

Otto, W., & Askov, E. The Wisconsin design for reading skill development: Rationale and guidelines, II: Components of the system. Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin. June 1970.

Warburton, F. W., & Southgate, V. *I.t.a.: An independent evaluation*. London, Newgate Press, 1969.

ROBERT L. AARON
University of Georgia

# DESIGN CONCEPTS FOR CONTINGENCY MANAGEMENT OF DELINQUENT ADOLESCENTS

Curriculum work with the institutionalized delinquent adolescent is a growing field, coming of age with the advent of treatment concepts, institutional designs, and a more humanistic philosophy conceived to cope with the burgeoning rate of youthful offenders.

Contingency management of pupils with atypical learning problems has been introdced as a promising way to get quick, positive changes in behavior. Staats (1970), in his work with severely disabled adolescent readers who were also delinquent, showed that a program of contingencies based on improved low probability performance would produce a very significant increase in the frequency and rate of reading performance.

A key problem in work with delinquent adolescents is to help them bring their antisocial behaviors under control and establish more socially acceptable behaviors. Attending, responding, and following directions in a socially acceptable way (Hewett, 1970) are the first order of business in any juvenile correction center.

There are other behaviors that are obviously debilitating the performance of the delinquent in the "straight" society. These pupils do not readily respond to social praise and, as a result, do not sustain control under the non-reinforcing conditions of such praise. However, the proper adult model (that member of the staff who functions as a "high probability") can teach the value of social praise by developing an affect relationship with the pupil in which the concrete payoff for some success at reading performance can be accompanied by social praise. In turn, success in other aspects of reading performance can be rewarded by time with the high probability model. A simple point, but a very natural one that we do not properly recognize in our public school program.

Attitude is certainly a big factor in the failure pattern of the delinquent. Notable is his attitude about reading, test taking, schools and teachers.

Staats (1966) has shown that paying delinquent adolescents to take standardized tests results in a significant gain in reading achievement scores. It seems that the teacher's prior experiences with test giving cause her to operate in a way that is so neutral, or negatively reinforcing, to the child that failure is foreordained for the acting-out delinquent. *It is clear* that the delinquent sees no payoff in test taking and is not easily motivated to continue an endless string of negative experiences.

Attitudes are taught and reinforced by the experiences undergone by the pupil. In effect, the schools, teachers, and reading classes so condition the pupil's responses to negative experiences with them that all but the more dramatic programs, designed to readjust attitudes, are foredoomed to fail. It often happens that the more traditional institutional setting is structured to pay off for work in the laundry or the kitchen, and not for academic excellence, thus perpetuating the negative response to learning. In effect, the institution does not rehabilitate the very critical attitudes about school success, but seems bent on habituating the negative reaction to school by improperly managing the contingencies.

A most significant point in working with these pupils is to base the design on "contracts" that are small enough to be met by the teacher. In many institutions the emphasis on a complete reading program, and curriculum, is so stress-producing for the untrained teachers that they resort to the more classic school patterns and immediately are identified in the same negatively reinforcing way as the delinquent's public school teachers. Research writings, and more practical experience, clearly indicates that this field is laden with affect skill problems. If the teacher learns to begin with small cognitive contracts, while focusing on the affect, then a much higher level of reading success can be predicted.

When a juvenile is first admitted to a state center a natural reaction is to refuse to make a normal effort in the testing and academic programs. Because many centers have natural payoffs available which they tend to give away—games, special trips, movies, reading materials, music, etc.—a program of contingencies was instituted to get maximum achievement on tests very quickly, and willing participation in the achievement program.

Tests of sight vocabulary, flashed and untimed, were administered under alternate reward and non-reward conditions. Flashed recall was worth three points, remembering letter order two points, and untimed analysis one point. Under all three conditions the students, when paid, showed a significantly superior performance level to the non-rewarded pupils' performance during the testing. Since test-taking has been such a negatively reinforcing experience this program of teaching the value of test-taking is continued until the pupil can begin to succeed on simple tests, which then allows the instructor to reinforce with knowledge of results, social praise, and social recognition. The ultimate objective is to send the pupil back to the public school system with a willingness to give maximum effort on the tests required by the school.

It is imperative in both the short and long-term institutional settings to quickly get the pupil involved in an achievement activity. The public school program has proven negative and unrewarding to the pupil, so the emphasis must be shifted to some payoff that will be immediately effective. In order to combine the effect of the contingencies and peer reinforcement, which came to be one key thing the newly arrived pupils were unable to do without, a combination of group payoffs and weekly cash-ins was instituted. It was immediately apparent that pupils who did not respond to the academic program, at first, readily came around and performed after finding themselves confined on Friday while their peers were cashing in their points for music, games, comics, books, magazines, and other activities. Once weekly, the pupils were assembled to get their tokens for the achievement activities during the week. This became a period when the peer group delivered a very high level of social praise as the totals were counted out for each person. This had its effect on the non-performing members and stimulated them to more effort.

A program of reading in high-interest pocketbooks was begun with an emphasis on reports being made on these books. The pupils readily purchased the books with their tokens but languished in their reading of them, and were averaging only one report every two weeks. A design was instituted in which the comprehension skills were classified as literal or interpretive, as in Bloom's Taxonomy, and points assigned to the skills in each area, with the interpretive skills worth more than the literal. Pupils could answer the questions in writing or orally. Immediately the rate of reports increased to five every two weeks, and reports were now coming from pupils who had never given a report before. By further addition of tape recording of the reports as another way of meeting the contract, an additional increase was achieved.

This simple design, interplaying intrinsic and extrinsic motivation, kept the stress on the original purpose—to get more book reading behavior on the part of the pupils, in hopes that they will take back to the public schools a positive response to reading.

A subjective, performance-based testing program was instituted, as in any remedial diagnostic program, because it is much easier to use the results to define small contracts and logical beginning points for work with the pupils. A key factor was the need for a precise curriculum entry point where work could begin immediately, and on a sufficiently low level to allow a maximum amount of success. Tests of contractions, inflected endings, phonics sounds, and other similar ones covering the basic terminal skills in reading, were used in a successive approximation strategy. Once a test was failed the pupil was placed in that skill group for remediation. The critical consideration in the concept is the fact that this straightforward design is highly reinforcing to the teacher. any current attempts at coningency managemen of learning problems are implementing designs that involve too may, or too big, contracts for the pupil and, most important, for the teacher. A close look at several nearby behavior modification projects that have failed reveals that new curriculum materials, a new classroom design, inadequate skills mastery by teachers, an inadequate or inappropriate testing program, covering too many curriculum subjects, and instituting the components all at once is much too big a contract for the teacher. That's the rub! We jazz it up for the pupils. Individualize!, we tell the teacher. Meet him where he is learning! Reinforce! But, the whole scheme, so glibly verbalized, is too much at once for the teacher and becomes a very negative experience as the pressure mounts from the need to fit all the components into a neat, experimental design. The real truth is that we do not understand the contingency management concept because we *do not* recognize that we must also make small contracts with the teacher so we can get success from her, reinforce the success, then build on this sane, humanistic approach.

For the delinquent adolescents who are disabled in reading it is important to find an adult model who is reinforcing to them, as soon as possible after they arrive at the center. If the teacher is compromised by a necessity to implement too many of the components at once, then she will invariably be cast in the role of schoolteacher, a negatively reinforcing experience for the juvenile. As a result, the resistance the teacher meets from the pupils will manage her in the direction of more and more negative, punitive reactions to the pupil. It is vital that the teacher who works with the delinquent adolescent's reading problems learn to make small contracts for the pupil. It is equally important that the system not define contracts for the teacher which are so big that she will pass on to the pupils the same old impossible load and negative reinforcement, thus creating an affect barrier between herself and the pupils.

Diagnostic teaching is a catch word we see much of in today's reading education literature. If the behavioral goals, mastery of teaching skills, and affect are there, the addition of small steps and a concrete reward system for teachers can make this catch word a reality. Only the most capable of teachers can properly assess and store feedback on pupil skill mastery in her mind, then react with effective remediation. If small steps and a concrete token system are combined, then the teacher can effectively reinforce herself about her successes and failures in the day-to-day teaching activities. The concrete act of giving tokens and seeing tokens earned should immediately let the teacher tabulate data on the pupils with which she did not succeed on a particular day. A color chip scheme can also serve as a means of letting her define the precise skills on which particular pupils did not succeed. In effect, she will be using a variation of Lindsley's concept of self-feedback.

Almost all pupils with reading problems, especially the delinquent adolescent, are quite weak in sight vocabulary. Flashed presentation of a wide range of sight words clearly shows that they can often name, in correct left to right order, all the letters in the word but cannot call the word, and have no apparent perceptual strategies of special letters, letter shapes, or letter-sound order by which to hold onto memory for the word. A program of strategy development involving interaction between the pupil and teacher, and a contingency management program with payoffs for successive steps mastered leading to a successful visual-perceptual strategy, should produce in the pupils a processing style that will make them more independent, and more efficient, in their word-recognition behavior.

A second program of sight word improvement grew out of an observation about the large number of difficult sight words that appear in modern comic books. A survey of one of the Archie comics resulted in a list of 104 sight words not known by the population of boys at one of the centers. Many of these were especially interesting because they were words like "truth, honesty, selfish, love, flesh, indecent," etc., that have therapeutic implications for these pupils.

Pupils earn points for each word learned as a flashed sight and meaning word. These words are then written into stories with readability levels to suit the instructional level of the pupil. These serve as reinforcement activities and as a check on the relationship between this vocabulary learning program and reading performance improvement.

The trend in delinquent adolescents is for the rate of girls to zoom and for all institutions to enroll a population with a much higher reading level. The coming trend in most states is for the number of institutionalized female delinquents to exced that of males.

In order to help attack what is apaprently deficit in critical thinking behavior a program using materials such as the Turner-Livingston Curriculum Motivated Series was developed. Many of the stories are very close to the experiences of the pupils and are used as a surrogate on which the pupils may try out their emotions and on which to practice the critical thinking skills before the teacher leads them through the "what would you do?" step, and into group interaction involving their own problems.

Taken as a total the previous suggestions could constitute a quite complete beginning curriculum for the juvenile detention centers, and, with some variations, would add significantly to the public school curriculum for the deviant adolescent.

## REFERENCES

Hewett, F. The Santa Monica project: demonstration and evaluation of the engineered classroom — phase II. ERIC ED038809, 1970.

Staats, A. Cognitive behavior modification: motivated learning reading treatment with subprofessional therapy technicians. *Journal of Behavior Research and Therapy*, 1967, 5, 283-294.

Staats, A., & Carlson, C. Classical conditioning of emotional responses (meaning, attitudes, values, interests) and effects on social behavior: a bibliography. Honolulu: Hawaii University, April 1970.

JOAN S. GRICE
South Colonie Schools, New York

LEE R. WOLFE
New York State Education Department

# THE DIFFERENTIAL EFFECTS OF PEER CORRECTION VS. TEACHER CORRECTION OF CLASSWORK ON SELECTED READING CRITERIA

While the subjects of this expericental study were first grade students, the general design of the experiment and its theoretical rationale seem sufficiently generalizable for inclusion in the present yearbook. The desire to maintain a positive image before one's peers, at any age, is a universal human trait for normal, healthy persons. One will not find many differences, except for age, in the behavior of reading students in Adult Continuing Education programs and these subjects.

Child development psychologists have observed that a child's desire for a favorable peer image is an important determinant of his social behavior and academic striving in the classroom. A measure of the strength of this desire for peer acceptance is the child's striving to conform to the traits of the dominant reference group. One of the variables within the classroom involving the cause-effect relationship between peer attitudes and behavior which may be used advantageously to increase school related learning is the knowledge that his peers possess the capability to monitor his daily achievement through the process of correcting his daily classroom assignments or seeking such information from a friend who has corrected the paper of a student of special interest. Primary age children tend to be secretive about their low grades and open about their high grades as a method of maintaining and improving their peer standing.

Behavioral psychologists have found knowledge of results to be a positive reinforcement which is beneficial to learning, especially when the results are favorable to the learner. The usual experimental format involves either the student-self or student-teacher relationship rather than the student-peer relationship. It is of considerable interest and importance to determine the reinforcement effects of knowledge of results fed back to the student by a peer rather than a teacher. What would the effects of knowing that a student's peers possess immediate knowledge of his achievement, even prior to himself, be on his academic striving and more specifically, his reading achievement? The experiment presented herein was designed to study the differential effects on the reading achievement of first grade students through the device of peer correction of classwork assignments.

Many teachers feel that too much of their time is spent in the correction of papers. The teacher could use her teaching ability and time more effectively if this could be done by other means with no adverse affects on the child's learning achievements. She could handle more children in the classroom and/or she would have more time for teaching on an individual or small group basis. Peer correction, use of paraprofessionals, and/or helping mothers could be used in this capacity. This study focuses on the alternative of using students to correct classwork papers.

Peer correction of classwork is a learning experience for the child doing the correcting. It may reinforce learning and make him aware of the importance of accuracy and neatness. When the low ability group corrects papers from the high ability group attention will be focused on exposure to new words and a challenege may be present to understand work beyond their own level of achievement. In addition, peer corrected papers may be checked more carefully for mistakes in correction when they are returned to the students. Thus the learning experience of checking one's own work is achieved and provides reinforcement for the particular learning activity.

An important determinant of behavior concerns the groups to which an individual belongs or with which he identifies himself. Striving for a better or more secure peer image creates a stronger learning set for the individual child. The literature surveyed discusses the ways educators have made use of this fact in other countries; research findings of others on ego involvement and self-concept; and the significance effects of peer relationships on learning.

### *Peer Pressure in Other Countries*

The system of collective education followed in the kibbutzim in Israel provides a unique opoprtunity to study the effects of a pattern of socialization that stresses group living and the importance of peer relationships from the very beginning of one's life. The ties between the children who are raised together are very strong, but they do not approach the intensity of relations between siblings brought up in a normal family atmosphere; however, they are much stronger than that of friendship. Competition exists between groups rather than within groups. The children tend to control each other's behavior and there is a strong tendency to conform to group standards according to Faigin (1958).

The Soviet Union has given much formal attention to the shaping of conforming personal behavior. They systematically use group pressure to train their youth according to Brofenbrenner (1962), the primary objective of Communist education is the development of socialist morality rather than the learning of subject matter. Anton Semyonivich Makarenko, a leader in Russian education, takes the viewpoint that the parent's authority over his child is delegated to him by the State and when there is a conflict betwen the values of the family and society, society has the priority.

Marenko's basic thesis is that personality development can best occur through productive activity in a social "collective" (an organized group engaged in some socially useful enterprise). The concept of peer importance and competition is stressed throughout a Rusian child's education. His contribution to group achievement is highly rewarded. He is pressured to help other members and to show appreciation for their contributions to the collective. The grade that each person receives depends on the overall performance of his unit, thus, good students find it to their advantage to help the poor ones.

Several other principles of Communist character education illustrate the importance of peer involvement. As soon as possible, the job of evaluating the behavior of individuals and of giving rewards is delegated to the members of the collective. Competition between groups is used as the prinicpal means for motivating achievement. Each member of the collective is encouraged to observe deviant behavior by his fellows and is given the opportunity to report his observation to the group. Telling on one's peers is esteemed and rewarded as civic duty.

### *Ego Involvement in Learning*

The meaning which a task has for the individual must be considered from the standpoint of its relationship to his self-concept. Generally, the individual's tendency to strive to make a good showing is guaranteed by his need to maintain status and to derive from the situation in which the student finds himself some degree of social approval from persons whose appraisal he values (Sears, 1970).

A change in motivation was found by Davids and White (1958) to occur when there was ego involvement. It was shown that disturbed children's goal setting after failure was most affected by: the nature of the task; the threat of loss of self-esteem; and the threat of loss of the experimenters' admiration.

Definite conclusions about the effects of anxiety on motivation still remain to be determined. A study by Reed (1960) concluded that both severe and very low levels of anxiety tend to depress learning while mild anxiety may function positively for some kinds of learning, the variations depending on differences in method. Sarason (1960) proposed that when induced anxiety produces responses that are compatible with the learning task, performance is facilitated. One researcher, Lighthall (1964), believes that anxiety exerts only an inhibiting effect on problems-olving effectiveness, while fear can exert a positive influence. He attributes this to the fact that the nature of fear is a conscious reaction to the situation.

A poor self-concept implies a lack of confidence that adversely affects a child's general adjustment to his environment. A study by Roth (1959) hypothesized that if demands for reading improvement are viewed as a personal threat to an individual, he will defend himself against the threat in order to maintain his self-concept. His defenses against the threat include distortion of the experience or withdrawal from the situation. Roth's hypothesis was partially confirmed.

For several years, Sears and Sherman (1964) have been engaged in a large-scale research program using behavior observations and other measures to study the achievement and self-esteem of elementary school children in their own classrooms. They found a significant correlation between general academic achievement and popularity.

Two investigators reported significant correlations between peer acceptance and self-concept scores. A study by Reese (1962) found acceptance and self-concept scores. A study by Reese (1962) found acceptance of others and popularity to be related to self-concept measures.

*The Significance of Peer Relationships*

Studies indicate a significant positive relationship between social acceptance and achievement, whether the study concerned elementary, high school or college populations. In a study by Williams (1961) achievement was significantly related to acceptance extended *to* the group rather than acceptance *by* the group. Hall and Gaeddert (1960) reported a moderately high relationship between friendship rating scales and grade averages of college students in fraternities and sororities. In a study by Elkins (1958) of characteristics associated with sociometric choice status, a slight tendency was found for highly chosen children to be more intelligent and to achieve higher academically. A significant relationship between peer status and reading achievement at every socioeconomic level was found by Portfield and Schlichting (1961). Arithmetic ability was found by Hudgins, Smith and Johnson (1962) to be significantly correlated to peer acceptance.

A study conducted by Patterson and Anderson (1964) tested the effects of peres as reinforcing agents using the marble-in-the-hole game.

Children in second and third grades showed significantly greater responsiveness to friends as reinforcing agents than non-friends.

A experimental approach was used in a study on the effect of value symbols and competition on the problem-solving behavior of five-year olds. The experiment revealed that performance varied in relationship to knowledge of their opopnents' performance. The results of this study by Leibowitz (1966) leave in question the issue of whether or not the children perceived the peer as an ally or an opponent. Studies of American children generally confirm Piaget's observations that "a child's morality changes as he grows older, from strict and specific moral rules deriving force from parental authority, to more general principles supported by groups of equals." (Quoted by Mac Rae, 1954). Peer group influences are extremely effective during the school years.

Twenty years ago, a study by Ausubel (1951) reported about "prestige motivation" in which children tended to work more efficiently and more quickly when the results of their work were made known to the group than when the results were known only by the individual.

Experimental evidence is not consistent: there is apparent evidence that anxiety can either inhibit or facilitate intellectual performance. It appears that many of the studies claiming to investigate anxiey are really studies of fear. The close connection between repression and anxiety seems to have ben overlooked. Depending on one's response to fear (approach or avoidance) it could exert an inhibiting or a facilitating effect on learning and performance.

The idea of a direct and fairly simple relationship between drive, as manifested by anxiety, and learning is seldom encountered in the research literature. It is reasonable to suppose that the level of a person's achievement is related to his drive state. If the drive state is influenced by a threat to one's peer image, it should be productive to learn what effect this would have on a specific learning situation.

## THE EXPERIMENT

The subjects for study of the differential effects of peer correction vs. teacher correction of classwork on subsequent reading achievement criteria were first grade students in a suburban elementary school. The students were divided into an experimental group and a control group. Each member of the control group had a pair-mate equivalent in the experimental group. The pairs were matched on: reading ability based on scores from the regular reading program; Metropolitan Test scores; and sex. Since class membership was 24 students, a total of 12 matched pairs were formed. Before the experiment, the class was organized into three ability levels; however, for purposes of this experiment, the "average" aiblity students were allocated to high and low ability groups to provide for two levels of ability within the experiment.

The regular classroom routines were maintained throughout the experiment period. No significant changes were made in the instructor, methods of instruction or time and place of instruction. In this way, major sources of variance were controlled.

The control group continued to hand in their classwork for teacher correction. No changes were made in the manner in which this had been done prior to the experiment. The teacher corrected and returned the reading worksheets to the control group students as usual. The experimental group had their papers corrected by other students in the experimental group. In the course of the experiment all experimental group students corrected several papers for each of the other students in the experimental group. In this way, any special relatonships between peers were randomized.

The duration of this peer correction experiment was three weeks. During this period, peer correction activities were completed on approximately 40 reading worksheets for each child in the experimental group.

In order to reduce the Hawthorne effect, i.e., experimental groups trying harder to succeed in order to help the experimenter and/or the control group's extra effort to keep up with the experimental group, the students were not informed that an experiment was being conducted. They were not told when it started, finished or what the objectives of the experiment were. The teacher simply announced that she wished to see how well students can be relied on to correct their friends' papers honestly and accurately. This focused any Hawthorne effect on reliable and accurate correcting and away from the main effects of the experiment under study. In this way, the experimental design took advantage of the Hawthorne phenomena and avoided its negative effects.

The experimental group was told that their reading worksheets would be corrected by their peers in the experimental group "until Easter vacation" and the control group would have similar treatment "after Easter vacation." This approach was taken in order to reduce two kinds of anxiety which may have been aroused in the students not selected for peer correction first; the "wonder" why they were not chosen; and the "threat to peer image" in the entire classroom during the treatment period.

The specific hypotheses under study were:

1) If children in the experimental group correct each other's papers, then the control group and the experimental group will differ significantly in *word recognition*, as measured by the Primary Reading Profiles, Level I Test, after the treatment period and will favor the experimental group.

2) If children in the experimental group correct each other's papers, then the control group and the experimental group will differ significantly in *word attack skills*, as measured by the Primary Reading Profiles, Level I Test, after the treatment period and will favor the experimental group.

3) If children in the experimental group correct each other's papers, then the control group and the experimental group will differ significantly in *reading comprehension*, as measured by the Primary Reading Profiles, Level I Test, and will favor the experimental group.

4) If children in the experimental group correct each other's papers, then the low ability experimental group will show a more significant increase in all three reading skills than the high ability experimental group.

## RESULTS

Table 1 presents the analysis of the *Primary Reading Profiles, Level I Test*, given at the conclusion of the treatment for all pairs on April 8, 1971. Table 2 presents the analysis of composite scores on the *Primary Reading Profiles Test* by ability grouping.

The discussion of hypotheses 1, 2, and 3 follows from Table 1 and the discussion of hypothesis 4 from Table 1 and Table 2.

The first hypothesis of the study; if children in the experimental group correct each other's papers, then the control group and the experimental group will differ significantly in *word recognition* after the treatment period and will favor the experimental group was not supported by these data at the $p < .05$ level. However, the experimental group's mean (39.17) was higher than the control group's mean (35.67) after the treatment, and this difference is significant at the alpha level of .10. Since t distribtuions tend to be leptokurtic for small sample sizes acceptance levels for significance are often reduced to .10 if the results are interpreted with appropriate caution. The t of 1.46 is well aobve the value $t = 0.697$ needed for significance at the .10 level for directional hypotheses. Therefore, the first hypothesis is retained with the understanding that these results could occur, by chance, one in every ten replications of this design.

The second hypothesis of the study; if children in the experimental group correct each other's papers, then the control group and the experimental group will differ significantly in *word attack skills* after the treatment period and will favor the experimental group was supported ($t = 1.93$) by these data and is therefore retained. Word attack skills differed significantly between the two groups and the results favored the students in the experimental group. The treatment effect of peer correction in this study evidently does increase achievement in the area of word attack skills.

The third hypothesis of the study; if children in the experimental group correct each other's papers, then the control group and the

Table 1

Means, Standard Deviations and t Tests for Primary Reading Profile Criterion Variables for All Pairs (N = 12)

| Primary Reading Profile Sub-tests | Experimental | | Control | | t Value |
|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | |
| H-1 Word Recognition | 39.17 | 9.18 | 35.67 | 10.54 | 1.46* |
| H-2 Word Attack | 14.0 | 4.62 | 11.58 | 4.86 | 1.93** |
| H-3 Reading Comprehension | 31.17 | 10.58 | 29.75 | 13.74 | 0.45 |
| H-4 Composition Scores[a] | 84.33 | 20.96 | 76.92 | 27.03 | 2.79** |

* $p < .10$
** $p < .05$
[a] Composit scores derived from scores on Word Recognition, Word Attack and Reading Comprehension.

experimental group will differ significantly in *reading comprehension* and will favor the experimental group, was not supported (t = 0.45) by these data and therefore may not be retained.

The fourth hypothesis of the study; if children in the experimental group correct each other's papers, then the low ability experimental group will show a more significant increase in all three reading skills than the high ability experimental group was supported by those data from Table 2 and is therefore retained. There was no significant difference (t = 0.78) between the composite scores of the high ability control group and the high ability experimental group. However, there was a significant difference (t = 2.36) between the composite

Table 2

Means, Standard Deviations and t Tests for Composite Scores on the Primary Reading Profile Test by Ability Grouping

| Level | Mean | S.D. | t Value |
|---|---|---|---|
| *High:* (N = 6) | | | |
| Experimental | 102.83 | 5.98 | 0.78 |
| Control | 100.50 | 11.81 | |
| *Low* (N = 6) | | | |
| Experimental | 65.83 | 12.58 | 2.36* |
| Control | 53.33 | 14.48 | |

* $p < .05$

scores of the two low ability groups in favor of the experimental group. Evidently the students in the low ability experimental group in this study were more affected by the treatment of peer-correction of their school work and their reading skills increased accordingly.

## CONCLUSIONS AND SUMMARY

Three of the four hypotheses in the study were accepted. It can, therefore, be assumed that the difference in the performance of the control group and the experimental group was due to the treatment, peer correction of school work, within the experimental group.

The third hypothesis which concerned an increase in reading comprehension by the experimental group after treatment was rejected by these data. Perhaps one of the reasons for this rejection was the shortness of the treatment period.

The results of this study imply that peer correction of classwork does increase a child's reading abilities, specifically his word recognition and word attack skills. Reading comprehension is a more complicated process and was not affected by the treatment within the treatment period. The need to maintain peer status in the classroom evidently is a strong motivating factor for achievement.

Future research is needed before attempting to establish generalizations from these findings applicable to other groups of school children. The effect of a longer treatment period should be studied in regard to reading comprehension. Research also is needed to discover if one can generalize from these results into other areas of the elementary school curriculum.

The purpose of this experiment was to study the effects of peer correction and knowledge of school work (reading worksheets) in a first grade classroom and to discover any relationships between this and an increase in reading skills. Only half of the children took part in the experiment. On a standaradized reading test, they scored significantly higher in word recognition and word attack skills after a three week treatment period. The children of "lower ability" made more significant increases in all reading skills than those children of "higher ability' which may indicate their greater concern and need to protect their peer image.

## REFERENCES

Ausubel, D. Prestige motivation of gifted children. *Genetic Psychology Monograph*, 1951, *43*, 53-117.

Bronfenbrenner, U. Soviet methods of character education: some implications for research. *Religious Education*, July-August 1962.

Davids, A., & White A. A. Effects of success, failure and social facilitation on level of aspiration in emotionally disturbed and normal children. *Journal of Personality*, 1958, *26*, 77-93.

Elkins, D. Some factors related to the choice-status of ninety-eight grade children in a school society. *Genetic Psychology Monograph*, 1958, *58*, 207-272.

Faigin, H. Social behavior of young children int he kibbutz. *Journal of Abnormal and Social Psychology,* 1958, *56,* 117-129.

Hall, W. E., & Gaeddert, W. Social skills and their relationships to scholastic achievement. *Journal of Genetic Psychology,* 1960, *96,* 269-273.

Hudgins, B. B., Smith, L. M., & Johnson, T. J. The child's perception of his classmates. *Journal of Genetic Psychology,* 1962, *101,* 401-405.

Leibowitz, S. L. F. The motivational effect of value symbols and competition upon problem-solving behavior in children. *Journal of Genetic Psychology,* 1966, *108,* 327-332.

Lighthall, F. F. *Anxiety as related to thinking and forgetting.* Washington, D. C.: National Education Association, 1964.

MacRae, D., Jr. A test of Piaget's theories of moral development. *Journal Abnormal and Social Psychology,* 1954, *49,* 14-18.

Patterson, G. R., & Anderson, D. Peers as social reinforcers. *Child Development,* 1964, *35,* 951-960.

Porterfield, O. V., & Schlichting, H. F. Peer status and reading achievement. *Journal of American Research,* 1961, *54,* 291-297.

Porterfield, O. V., & Schlichting, H. F. *Primary reading profiles,* level I test. Boston: Houghton Mifflin, 1967.

Reed, H. B., Jr. Anxiety, the ambivalent variable. *Harvard Educational Review,* 1960, *30,* 141-153.

Reese, H. W. Relationships between self-acceptance and sociometric choices. *Journal of Abnormal and Social Psychology,* 1962, *62,* 666-672.

Roth, R. M. The role of self-concept in achievement. *Journal of Educational Psychology,* 1959, *27,* 132-135.

Sarason, S. B., et al. *Anxiety in elementary school children.* New York: Wiley, 1960.

Sears, P. S. Levels of aspiration in academically successful and unsuccessful children. *Abnormal Psychology,* 1940, *35,* 102-119.

Sears, P. S., & Sherman, V. S. *In pursuit of self-esteem.* : Wadsworth, 1964.

Williams, M. F. Acceptance and performance among gifted elementary school children. *Educational Research Bulletin,* 1961, *37,* 183-186.

CARL A. LEFEVRE
Temple University

# READING BY PATTERNS: A PSYCHOLINGUISTIC REMEDIAL TUTORIAL PROGRAM FOR YOUNG ADULTS

The original idea for the psycholinguistic approach embodied in *Reading by Patterns* developed as I came to realize the enomity of the problems, at all grade levels in our schools, variously called word calling, or word naming, or barking at the page. It struck me forcibly that students so afflicted have literally learned only what they have been painstakingly taught—nothing else: letters, sounds, syllables, words. Other children somehow catch the trick of reading, but these unfortunates do not. Instead, they regard the task of reading printed English as the laborious identification of word after word in isolation, as in a list printed from left to right. Once a word has been more or less identified—called, named, or barked—the student moves across the white space between the words and attacks the next word. The internalized silent reading of such readers no doubt mirrors this procedure.

As I have noted elsewhere in my books and articles (1964a, 1968a, 1970), this procedure at best fragments English sentences and destroys the normal rhythmic and intonational patterns that hold sentences together and mould them into whole meaning-bearing units. One of the most grievous outcomes of this rhyth-destroying procedure, however, is that many students never learn to sound out words in this way at all. Instead, they stumble across one white space after another, desperately guessing at the connections of letters, sounds, and syllables, often failing miserably, word after word. They are so preoccupied with the tiniest bits of information on the printed page that they cannot possibly attend to the larger segments and structures required for successful decoding of whole messages (Compare Smith and Holms, 1971; and Weaver, 1967). Such reading cripples are unable to apply what they have only partially perceived through unsuccessful rote learning. They cannot, in short, even identify words successfully. It is of small help to such students to be advised to read with expression, or exhort to develop sentence sense. Their need of systematic instruction in how to decode whole sentences so as to derive meaning from the relationships of the main sentence parts to the entire message is painfully clear. *They need to be taught how to develop sentence sense.*

This program is a beginning toward that end. The full title–*Reading by Patterns: A Programed Guide to Reading Sentences and Paragraphs* (Lefevre, Lefevre and Shore, 1971)–emphasizes that the most significant meaning-bearing patterns in English are well above the levels of letters, sounds, syllables, and words. The principal aim of the instruction is to help the student transform his normal fluency in speaking oral English into a parallel ability to read simple printed English sentences and paragraphs. *This is the essential trick of successful reading.* It is the first step on the road to reading more sophisticated printed English. The unique, distinguishing feature of this approach is that it helps students decode and recode sentences in terms of main sentence parts (not parts of speech). But this is not grammar instruction. No verbal definitions of any kind are required; instead, a minimum of terms is used, learned gradually by association, repetition, imitation, and application. Thus the student learns to associate the main structural parts with the message of each sentence. The thrust of the entire program is to develop comprehension of messages conveyed through printed English sentences. As the instruction progresses, successive units call attention to distinctive elements within main sentence parts, and to the relationships and interdependencies of sentences within paragraphs.

This particular portion of the projected program focuses upon the N V N sentence pattern: Noun part, Verb part, completer. These simple terms are used instead of Subject and Predicate; or Noun Phrase and Verb Phrases; or Subject, Verb, Object; because Noun part, Verb part, and completer are simpler and more functional. Moreover, the N V N pattern is the basic English sentence pattern, and the most commonly used; many other English sentence patterns also have three parts, and this parallelism with other patterns is reinforcing and transferable. The authors also believe that a student who knows that no new pattern will be "sprung" on him has confidence enabling him to focus all his attention on the essential points throughout this program.

This approach to reading instruction has a psycholinguistic base. Very simply, the term *psycholinguistics* incorporates overlapping areas between psychology and linguistics. The main hypothesis underlying this psycholinguistic rationale is as follows. *All thought that is distinctively human* (not shared by lower animals), *incorporating both cognitive and affective experience, is quite literally bonded to language: human thought and language* (which, in so far as it is a symbolic system rather than a system of signs, is a distinctively human attribute) *form an integrated whole. Distinctively human thought does not exist without distinctively human language.* I share this view with Sapir (1972), Fries (1945), Vigotsky (1961), and Pike (1964a and b), among others. Speaking, listening, writing, reading, and thinking, all are interrelated psycholinguistic processes. Reading English requires interpretation of printed English, which is ultimately based upon and derived from spoken English. In this sense, then, I define reading as a psycholinguistic process.

This program is also psycholinguistic in the sense that it offers disabled readers a fresh chance to come at reading, a method quite different from their previous attempts to attack letters, sounds, syllables, and words. By the time they reach the 7th grade, most ordinary students have been repeatedly exposed to a variety of word attack skills. This program offers them a chance to observe and practice the interrelationships among thinking, speaking, listening, reading, and writing. Finally, and most importantly, it offers them initial success and repeated success after years and years of reading failure.

## BRIEF DESCRIPTION OF THE PRESENT PROGRAM

In developing this program, I have been ably assisted by Helen E. Lefevre and David Shore; Wendell W. Weaver has served as consultant on testing. Following is summary information from the cover page of *Reading by Patterns:*

> *Capsule statement:*
>
> This edition of READING BY PATTERNS is about the fifth generation in the development of this material; it is perhaps two thirds the estimated length of the present program. Portions of this material in various stages have already been tested by more than five hundred students in the Philadelphia area and in Georgia. It is currently being used in one-to-one tutorial experiences by undergraduate and graduate students at Temple University. The course is entitled Language and Reading Problems of Young Adults; it is designed for college students training to teach pupils in grades 7-12 whose reading levels are from 3 to 5 as measured by such standard tests as the Iowa or Stanford.
>
> *Programed instruction:*
>
> A unit is the smallest self-contained portion of the program; each unit consists of 9 to 16 frames requiring a response by the student. A student should always complete a minimum of one

> unit per working session, but he may complete as many units as time and inclination allow.
>
> *Procedure:*
>
> The teacher, a well-trained tutor (or a good tape recording) reads the directions aloud to an individual student, who follows the directions visually as the teacher reads; since all parts of the program are read aloud to him, the student does not have to be able to read independently in order to perform the assigned tasks successfully. The visual "read along" experience is believed to be a major strength of this method. Each student uses a sheet of cardboard or manila paper to mask the frames he is not working on. As he completes each frame, he slips his mask down and exposes the next frame; if he chooses, he may go back to an earlier frame and repeat his work. The basic aim of every frame and every unit in this program is success for the student.

A word of caution. This linear program is conceived as a single strand in a comprehensive reading program; it is by no means intended to be a comprehensive reading program in itself. During the tutoring process, the simple language-comprehension learnings of this program may be applied directly to more complex language-comprehension problems in textbooks and other materials currently studied, as in social studies and science. Similarly, the language-comprehension elements may be less formally applied to motivating read-along experiences using high-interest materials for individual students. Students should also enjoy access and exposure to films, TV presentations, plays, and interesting oral readings either by professional readers, by the teacher, or by classmates. READING BY PATTERNS should be kept in perspective as a single strand of reading instruction having precise but limited objectives; it is not a comprehensive reading program. (See Lefevre 1964 and 1968 for comments on comprehensive programs.)

## A Sampling of Reading by Patterns

*Levels of reading difficully*

The first unit begins at the pre-primer level, yet this very first unit teaches the crucial concept that English sentence parts must be in the right order: Noun part first, Verb part second, completer third; English sentences must be written, read, and comprehended as unitary meaning-bearing patterns with the three main sentence parts in one, two, three order. The extreme simplicity of the first sentence has allowed every remedial reader who has tried the program to succeed in his very first effort. We feel that this initial success is of the utmost importance to students who have experienced nothing but reading failure for many years.

The level of difficulty progressively increases throughout the program, as the sentence parts become more complex, and as various elements of the more complex parts are sequentially called to the student's attention. To experience this program as the tutorial student experiences it, the reader of this article should provide himself with a mask, as described above under "Procedure"; and then advance deliberately from frame to frame, simulating the kind of concentrated attention that is required of the remedial reader. *At his level of reading competency, he is forced to think.*

The following examples include the complete revised worktext version of Unit 1; four selected frames from Unit 2; and the complete text of the concluding unit of this part of the program, Unit 54. These examples give a minimal representation required to illustrate the underlying psycholinguistic rationale of READING BY PATTERNS.

### *Unit One*

1. Three words are printed below. Read them to yourself while I read them aloud.

saw
Robert
I

No answer needed.

GO ON TO #2

2. Do the three words in item 1 make a sensible sentence as I read them to you? (answer yes or no)

Answer here

THE CORRECT ANSWER IS NO

3.

saw
Robert
I

These three words can be written as a sensible sentence. Can you do it? Write the sentence now on the line below

(1) I (2) SAW (3) ROBERT

The sentence, "I saw Robert" makes sense because the *sentence parts* are in the right order: 1, 2, 3; or first, second, third. That is what right order means.

4. Let's try a new sentence with the *sentence parts* printed below. Read them to yourself while I read them aloud.

*him*
(3. completer)

*called*
(2. verb part)

*The teacher*
(1. noun part)

No answer needed

**GO ON TO #5**

5. You can write a sensible sentence by writing the sentence parts in #4 in the right order. Start with the noun part, the ? . Write the noun part on the line below.

1. The

**THE TEACHER**

6.

| *him*<br>(3. completer) | *called*<br>(2. verb part) |
|---|---|

What sentence part would you add after *The teacher?* Answer below.

The teacher

**THE TEACHER CALLED**

7. What sentence part would complete the sentence?
The teacher called

**THE TEACHER CALLED HIM**

8. Read your answer to #7. Did you write the three sentence parts in the right order? (Yes or No)

**YES**

9. *The teacher called him* makes sense because the sentence parts are in the right order. The *noun part* comes: (a) first; (b) second; (c) third (choose one answer)

**FIRST**

10. *The teacher called him*. The *verb part* comes: (a) first; (b) second; (c) third

**SECOND**

11. *The teacher called him*. The *sentence completer* comes: (a) first; (b) second; (c) third

**THIRD**

12. In a sensible sentence, the sentence parts must be in the right ________?

**ORDER**

*Unit Two* (selected frames and commentary)

Unit 2 advances from Unit 1 in two principal ways: it introduces markedly longer sentence parts; and it introduces a frame in which the student must choose among three Noun parts, three Verb parts, and three completers.

Following are the first two frames of Unit 2:

1. In Unit 1 we made sentences by writing the three sentence parts in the right order. A new set of three sentence parts is printed below. Read them to yourself while I read them aloud.

*was watching*
(2. verb part)

*A very tall policeman*
(1. noun part)

*the long parade*
(3. completer)

No answer needed

---

GO ON TO #2

---

2. Let's begin our sentence with the *noun part* from #1.
Write it below.

---

A VERY TALL POLICEMAN

---

Frames 3, 4, and 5 follow the same procedure as in Unit 1.

Below is a new kind of problem requiring the tutorial student to make careful choices among nine sentence parts. Three sets of sentence parts are constructed so as to be error-proof, provided the student follows the right order of sentence parts. But because the student does not know that the problem is error-proof, he works diligently to make intelligent choices. He thinks.

6. A larger group of sentence parts is printed below. Read the sentence parts to yourself while I read them aloud.

---

*(1) Noun parts*
His brother
Your car
My motor bike

*(2) Verb parts*
hit
bumped
scratched

*(3) Completers*
the wall
my arm
his truck

---

No answer needed

---

7. You can write the sentence parts in #6 in the right order to make sensible sentences. Choose one noun part from #6 and write it on the *first* line in the box below.

---

(1) ______________________ (2) ______________________
(3) ______________________

---

HIS BROTHER *or* YOUR CAR *or* MY MOTOR BIKE

---

Frames 8, 9, 10, 11, 12 follow the same procedure as before, reinforcing the concept that to make sense, the three main parts of a sentence must flow together in the right order.

Units 3 through 54 deal with more involved sentence parts, and with successively more detailed attention to significant signals and markers within sentence parts. The first simple paragraph is presented in Unit 16. The first fifty three units all deal with various elements of sentence structure as they relate to communicating the total messages conveyed by simple sentences and paragraphs. Unit 54, however, deals entirely with comprehension of meaning, with no attention to structure whatsoever.

*Unit Fifty-four*

Unit 54 is the final unit of READING BY PATTERNS in its present form; it is the culmination of nine units dealing with various aspects of a single paragraph. It is doubtful that even a skilled reader at this reading level could answer these questions successfully without having worked with some, possibly all, of the eight sequential units leading up to it.

1. The paragraph that we began working with is reprinted below. Read it to yourself again while I read it aloud. Use it to help answer the questions that follow.
   1) A burning tropical sun baked the flat, empty, arid plain. 2) Many different animals surrounded the quiet shaded water hole. 3) A brace of powerful lions approached the inviting oasis. 4) Both big cats began drinking its cool water. 5) Several giant water buffaloes joined these great cats. 6) Neither thirsty lion seemed to notice these heavy-horned new neighbors. 7) No other wildlife joined the dangerout animals at the pond. 8) Most other species try to avoid mature lions. 9) Few animals will approach any full grown water buffaloes.
2. From reading this paragraph you know that around the waterhole there were: (a) only a few animals; (b) a lot of different kinds of animals; (c) some men and some animals.

   A LOT OF DIFFERENT KINDS OF ANIMALS
3. Which two types of the animals around the waterhole were the *most* dangerous?

   LIONS AND WATER BUFFALOES
4. The different animals came to the oasis: (a) to meet the other animals; (b) to get a cooling drink; (c) to find food.

   TO GET A COOLING DRINK
5. The oasis is probably located: (a) in a cold damp jungle; (b) in a hot dry desert; (c) in a wet steaming swamp.

   IN A HOT DRY DESERT
6. The oasis is probably: (a) hotter than the rest of the desert; (b) the same temperature as the rest of the desert; (c) cooler than the rest of the desert.

   COOLER THAN THE REST OF THE DESERT
7. The water buffaloes that approached the waterhole were probably: (a) smaller than most animals; (b) as large as most animals; (c) much larger than most animals.

   MUCH LARGER THAN MOST ANIMALS
8. The water buffaloes were much larger than most animals because they are described as ?

   GIANT
9. Most of the animals didn't approach the waterhole while the lions were there. Most of the animals probably: (a) weren't thirsty; (b) were tired; (c) were afraid of the lions.

WERE AFRAID OF THE LIONS

10. One species of animal that is not afraid of lions is the ?

WATER BUFFALO

11. We know that the water buffaloes were not afraid of the lions because: (a) they approached the lions; (b) they avoided lions; (c) they fought the lions.

THEY APPROACHED THE LIONS

12. From reading this paragraph, you know that: (a) all of the animals drank together; (b) the most dangerous animals drank first; (c) the fastest animals drank first.

THE MOST DANGEROUS ANIMALS DRANK FIRST

READING BY PATTERNS has been successfully used by remedial readers who speak both standard and non-standard American English dialects. Their ability to answer correctly the relatively difficulty questions in Unit 54 is illuminated by an important experimental finding:

> the use of pattern practice techniques can sharpen students' awareness of structural matters in such a way as to improve their comprehension of material that they read (Lin, 1965).

READING BY PATTERNS applies this insight directly and systematically to developing comprehension of sentences and paragraphs.

### REFERENCES

Fries, C. C. *The teaching and learning of english as a second language*. Ann Arbor: University of Michigan Press, 1945. (Pp. 57-61, Contextual orientation.)

Lefevre, C. A. *Linguistics and the teaching of reading*. New York: McGraw-Hill, 1964. (a)

Lefevre, C. A. A comprehensive linguistic view of reading. *Elementary English*, Oct. 1964, *41*, 43-50. (b)

Lefevre, C. A. The simplistic standard word-perception theory of reading. *Elementary English*, March 1968, *45*, 349-353. (a)

Lefevre, C. A. A multidisciplinary approach to reading: some projections. In K. S. Goodman (Ed.), *The psycholinguistic nature of the reading process*. Detroit: Wayne State University Press, 1968. Pp. 289-213, 333-336. (b)

Lefevre, C. A. *Linguistics, english, and the language arts*. Boston: Allyn and Bacon, 1970.

Lefevre, C. A., Lefevre, H. E., & Shore, D. *Reading by patterns: a programmed guide to reading sentences and paragraphs*. Philadelphia: Lingua Press, 1971.

Lin, S. C. *Pattern practice in the teaching of standard english to students with a non-standard dialect*. New York: Teachers College Press: 1965.

Pike, K. L. A linguistic contribution to composition: a hypothesis. *College Composition and Communication*, May 1964, *15*, 83-88. (a)

Pike, K. L. *Language in relation to a unified theory of the structure of human behavior*. New York: Humanities Press, 1964. (b)

Sapir, E. A. *Language: an introduction to the study of speech*. New York: Harcourt Brace, 1921.

Smith, F., & Holmes, D. L. The independence of letter, word, and meaning identification in reading. *Reading Research Quarterly*, Spring 1971, *6*, 394-415.

Vigotsky, L. S. Thought and speech. In S. Soporta (Ed.), *Psycholinguistics: a book of readings*. New York: Holt, Rinehart and Winston, 1961. Pp. 509-537.

Weaver, W. W. The word as the unit of language. *Journal of Reading*, Jan. 1967, *10*, 262-268.

PATRICIA R. McCLENDON
Winthrop College

# INDIVIDUALIZED INSTRUCTIONAL OBJECTIVES IN READING METHODS COURSES

Those involved in teacher education have long been concerned with the quality and effectiveness of programs in institutions of higher education. In 1967 the U. S. Office of Education funded ten institutions for the purpose of designing innovative program models (Model Elementary Teacher Education Project). Extensive descriptive volumes published from each participating university outlined possible procedures for individualization of instruction. Unfortunately, the suggested designs were unfeasible for institutions the size of Winthrop College which has a student body of 3,750, with 1,250 students enrolled in the School of Education. Several program models provided some direction in their stated underlying assumptions and the Dean of the School of Education urged staff members to utilize these in exploring alternative avenues for curriculum improvement.

## EVOLUTION OF THE PROGRAM

The decision to utilize individualized instructional objectives in reading methods courses was an outgrowth of planning for an overall teacher education program at Winthrop. Bases for curriculum planning embraced analysis of the:

characteristics of the society in today's world which have the potential for becoming even more prominent and perplexing in the future;

characteristics of the individual which seem to be vital if he is to function effectively in this society;

characteristics of the teacher who will be able to provide in his classroom, a micro-society which will nurture capable individuals;
envisioned roles of the elementary teacher as he provides an educational program from which will emerge such individuals; and
characteristics of a teacher education program designed to enable the elementary teacher to prepare for these roles.

Within this framework, a *Pilot Project for Individualizing Elementary Teacher Education* (McClendon, 1970) was instituted during the 1969-70 school year. A curriculum unit approach with individualized objectives served as the vehicle for individualizing instruction in the project which included only sixteen students and four courses.

Upon completion of the project, it became the task of various members of the School of Education staff to analyze the findings of the year's study in search of features which showed promise for an improved program, and to implement those features in the areas for which they were responsible. It was felt that many aspects of the Pilot Project could be incorporated into the area of reading instruction.

Winthrop College offers eight courses in the teaching of reading. One of these courses is a basic undergraduate study designed for elementary education majors. With the exception of a course in reading in the secondary school, which may be taken at either the undergraduate or graduate level, the remaining courses are primarily graduate level and directed toward the in-service teacher. Although individualized instructional objectives have been, or are in the process of being, developed for all courses, the following presentation relates to the basic undergraduate reading course for the elementary education major. This course presents the greatest challenge for individualizing instruction due to the number of students (approximately 150 each semester) and the diverse areas represented in the group—elementary, early childhood, and special education majors. At the request of the staff members in charge of the latter groups, separate sections are not set aside for either special education or early childhood education students.

### PURPOSE OF THE PROGRAM

As specific planning for the elementary reading course began, four purposes were given as much consideration as was course content.

1. To provide for the individual college student a more effective course of study which will foster many of the characteristics seen as needed in our present and emerging society.
2. To engage the undergraduate in the cooperative strategy for learning (rather than a receiving-telling mode) similar to that which he will be expected to implement with his own students.
3. To design an educational process which will be more likely to

provide the varying levels of competencies which will be needed as differentiated staffing in school faculties increases due, in part, to the existing teacher surplus.

4. To establish an instructional plan to counteract the existing situation in which a disproportionately high percentage of the instructor's time is devoted to the less capable student.

These four purposes remain prominent in continuing revisions of the program.

## DESCRIPTION OF THE PROGRAM

The plan presently in operation incorporates, in addition to individualized instructional objectives, team teaching and the extensive use of multi-media.

### *Organizational Format*

Three concerns of reading staff members were fundamental to the establishment of the organizational format being used.

A primary concern centered around the recognition that a relatively large number of students, in the first year of the program, found it difficult to function in a course which did not follow the familiar traditional pattern of "lecture-assign textbook readings-test-grade." Although the frustrations voiced by these students did not seem to have a detrimental effect in the cognitive domain (based on comparison of test scores), the apparent consequences in the area of the affective domain could not be ignored.

Secondly, sections of the reading courses, prior to the introduction of this program, ranged in class size from approximately 28 to 38 students. Such class size was not conducive to the degree of interaction which was desired.

The last concern pertained to the desire on the part of the reading staff to engage in team teaching.

Consequently, this course is organized so that each student attends one mass lecture (150 students) and two seminars (approximately 20 students) per week. One staff member has the major responsibility for the mass lectures, however the other two staff members present mass lectures in their areas of special competencies. It is hoped that eventually the mass lecture will be relegated to a position of minor importance if not discontinued entirely. Increased emphasis will be placed on individual study and seminars which should function as inquiry groups.

### *Types of Objectives*

Three types of objectives—required, test make-up, and elective—are employed in the program.

Perhaps the term "required objectives' 'appears incongruous in a plan which purports to be individualized, however, in this program they represent one facet of a competency-based structure. These objectives are listed in the Program Guide for the course and describe certain understandings and concepts which are considered basic for all students. The following, taken from the curriculum unit on word perception, is an example of the required objectives.

2:0405 After reading or hearing reports on selected research studies, the learner correctly identifies statements representing the findings.

Varied capacities for self-direction exhibited by previous students, combined with the large number of students enrolled in the course, seemed to mandate that the majority of the background information needed for required objectives be presented in mass lecture or in the main text used in the cours. The primary means of evaluation for these objetcives is four multiple-choice tests. The scores required on these tests depends on the level of competency set by the student as his goal.

Test make-up objectives are optional. Winthrop College has a system for grading which allows a student to take a course on an A-B-C-D-F basis or a S-U basis (S designating a grade of "C" or better). In an effort to minimize the potential negative aspects of grading, a procedure utilizing test make-up objectives was devised. Any student whose score falls below the level he has set for himself may analyze his test using an Answer Check Sheet (Fig. 1) and a Test Analysis Sheet (Fig. 2). He then chooses, from the questions he missed, those he considers most important and writes objectives specifying what needs to be done for clarification. Within one week he must submit a paper (based on his test make-up objectives) which reveals an understanding of the selected items and raises his score to the desired level.

Students setting as their goal an above average level of competency must complete a prescribed number of elective objectives in addition to maintaining a designated score on each test. Possible elective objectives are given in the Program Guide, however students are urged to modify these objectives and/or write their own elective objectives. Those objectives which are included in the Program Guide represent an attempt to have the student undertake a task similar to those required of a teacher in the classroom. For example, in the curriculum unit on word perception skills the following elective objective is listed.

2:0415 The learner, referring to characteristics of the "typical" disadvantaged child, outlines (in chart form) problems which this child might have in the development of word perception skills and suggestions for overcoming each problem. (No more than two pages)

In this objective the college student is expected to engage in the kind of thinking required of the classroom teacher as he plans his reading program. Suggestions for accomplishing each of these objectives are recorded on cassette tapes and made available for use on listening sta-

**NAME** **DATE** **SECTION**

| YOUR ANSWER* | CORRECT ANSWER | | | CORRECT ANSWER | YOUR ANSWER* |
|---|---|---|---|---|---|
| | C | 1 | 26 | E | |
| | E | 2 | 27 | B | |
| | D | 3 | 28 | C | |
| | A | 4 | 29 | E | |
| | C | 5 | 30 | A | |
| | B | 6 | 31 | D | |
| | B | 7 | 32 | D | |
| | D | 8 | 33 | C | |
| | A | 9 | 34 | A | |
| | E | 10 | 35 | E | |
| | C | 11 | 36 | A | |
| | A | 12 | 37 | C | |
| | E | 13 | 38 | B | |
| | B | 14 | 39 | E | |
| | D | 15 | 40 | D | |
| | B | 16 | 41 | A | |
| | A | 17 | 42 | A | |
| | A | 18 | 43 | D | |
| | D | 19 | 44 | C | |
| | B | 20 | 45 | D | |
| | A | 21 | 46 | E | |
| | E | 22 | 47 | A | |
| | C | 23 | 48 | B | |
| | C | 24 | 49 | E | |
| | E | 25 | 50 | D | |

*** If the answer you gave was *wrong, write the answer you gave* in this column.**

**Fig. 1. Answer check sheet**

tions located in the Materials Center. Each student undertaking elective objectives is expected to share his studies in seminars both through oral reports and contributions to discussions. He must evaluate his elective objectives in terms of an adaptation of *Bloom's Taxonomy*

## TEST ANALYSIS SHEET
## CURRICULUM UNITS 2:00 AND 2:01

Required Score ____
Your Score ________
No. of points
to be clarified _____

Name ________________ Date __________ Section ___

Numbers in parentheses indicate point value of question. Questions without parentheses have a one point value. *CIRCLE* THE QUESTIONS YOU MISSED AND *WRITE IN THE* ANSWERS *YOU* GAVE.

**Disadvantaged Child**
1_______(3)
2_______(3)
3_______(3)
4_______(3)
5_______(3)

**Historical Development**
6_______
7_______
8_______
9_______
10_______
11_______
12_______
13_______

**Devices/ Terms**
14_______
15_______
16_______
17_______
18_______
28_______
47_______(2)
48_______(2)
49_______(2)

**Physical Factors**
19_______(3)
20_______(3)

**Mental Factors**
21_______(3)
24_______(3)
29_______
30_______(3)
44_______(3)

**Motivational Factors**
25_______(2)
26_______(2)
27_______(2)

**Educational Factors**
22_______(3)
23_______(3)

**Readiness**
31_______(3)
32_______(3)
33_______(3)
34_______(3)
35_______(3)
36_______(3)
37_______(3)

**Persons**
38_______
39_______
40_______
41_______
42_______
43_______
45_______(3)
46_______(3)

50_______(2)

I will clarify questions ______________

PLANNED OBJECTIVES (S)

Fig. 2. Test Analysis Sheet

NAME_____________________________ SECTION______
MAJOR____________________
OBJECTIVE NUMBER__________ DATE____________
REASON_________________________________________
_______________________________________________
_______________________________________________

INSTRUCTOR'S EVALUATION

Inadequate at this level but adequate at a____________level
Excellent____________ Adequate____________
Unacceptable____________

Fig. 3 Contract Completion Form

*of Educational Objectives: Cognitive Domain* and designate the level of difficulty for each on a Contract Completion form, (Fig. 3).

Major differences exist between required and elective objectives particularly in the type of thinking required, the level of difficulty of th objectives according to *Bloom's Taxonomy of Educational Objectives,* the sources which are to be utilized, and the nature of the information. These differences are presented in Table 1.

Table 1
Major Difference Between Required and Elective Objectives

| **Differences** | **Required Objectives** | **Elective Objetcives** |
|---|---|---|
| **Type of Thinking Required** | **Convergent Thinking** | **Mainly Divergent Thinking** |
| **Level of Difficulty (Bloom's Taxonomy)** | **Predominantly 2.00 and 3.00 (Comprehension and Application)** | **Predominantly 3.00, 4.00, 5.00 and 6.00 (Application, Analysis, Synthesis, and Evaluation)** |
| **Sources for Information** | **Primarily mass lecture and the main text used (audio-visual materials are utilized in mass lectures and seminars)** | **Individual utilization of various library resources, audio and video tapes, monographs and pam- filmstrips and slides, curriculum guides, guide books, materials for reading instruction in the classroom.** |
| **Nature of Information** | **General concepts and understandings in the field of reading** | **Concepts and understandings from the field of reading related to the student's special area of concern** |

*Criteria for Various Levels of Competency*

The student who feels that he will be able to achieve at a superior level ("A") must attain, on three of four multiple-choice tests, a test score which is at least equal to 90 per cent of the highest score obtained on that test. Since any cut-off point is at best an arbitrary decision, a student may, in fact, fall a few points below the 90 per cent mark on several tests. Each test, except the final, must be brought up to the 90 per cent level or higher by test make-up objectives. In addition, the superior student must complete four elective objectives from specified curriculum units. At least three of his elective objectives must be at a 3.00 level of difficulty (Application) or higher.

Criteria for the above average level of achievement ("B") state that the student must attain test scores which are at least equal to 85 per cent of the highest score obtained on each test. If he scores below that point he must analyze his test and complete test make-up objectives to increase his score to 85 per cent or higher. Students functioning at this level may use test make-up objectives for each test (except the final) and still obtain a "B" in the course no matter how low his original test scores were. Three elective objectives are required of these students, with no specific level of difficulty designated.

To function at an average ("C") level, the student must show satisfactory mastery of the basic concepts and understandings of this area of study by achieving a score on each test which is at least equal to 75 per cent of the highest score attained. Test make-up objectives may be used to reach this requirement. No elective objectives must be submitted by this student.

## CONCLUSION

The program in progress at Winthrop College can best be described as emerging. Revisions are made each semester in light of student and staff evaluations of the course. As the staff becomes more competent in planning and implementing the program, and as additional multimedia materials are developed, the projected purposes of this design for individualizing instruction should be realized. Just as extensive planning was required prior to the inception of the program, future changes must also emanate from a sound and comprehensive rationale. It is much too easy for those who develop individualized instructional objectives to become skillful in stating, precisely, the trivial. No educational program can withstand such a miscarriage of learning.

## REFERENCES

McClendon, P. R. Pilot project for individualizing elementary teacher education—final report. Rock Hill, S. C.: Winthrop College, 1970. ERIC EDO40981.

WAYNE OTTO
University of Wisconsin

# RATIONAL AND OVERVIEW OF THE WISCONSIN DESIGN[1,2]

The Wisconsin Design for Reading Skill Development is based upon the assumption that reading can most effectively be taught by focusing upon the important subskills. A product of the Wisconsin Research and Development Center for Cognitive Learning, the Design is intended to assist teachers with the organization of instruction and the management of skill development in reading. Our purpose of this symposium is, first, to describe the Design to you, and second, to consider with you possibilities for its use in adult basic education programs.

## THE FRAMEWORK OF THE DESIGN

In developing the Design, we have made use of a framework that includes the following elements:

1. *Identification of Essential Content.* Agreement regarding the content essential to success in a given curriculum areas is necessary before there can be a straightforward approach to instruction. Efforts to specify essentials have not been very definitive except in the basic skill areas, where essential content can be reduced to essential skills.

---

[1] This is paper one (of five) from the symposium. A skill centered approach to reading instruction: The Wisconsin Design for Reading Skill Development.

[2] All the papers in this symposium were supported in part by the United States Office of Education. The opinions expressed here do not necessarily reflect the position or policy of the Office of Education and no official endorsement should be inferred.

In the area of reading we have identified "essential" skills that have consensual, historical and limited empirical support.

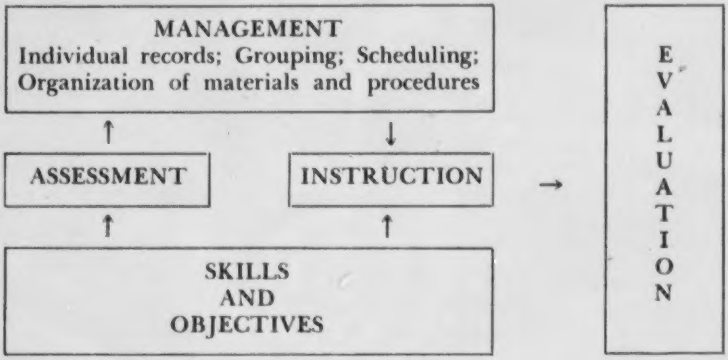
2. *Statement of Objectives.* Adequately stated objectives specify the criterion behaviors related to each essential skill. They specify the operational mastery levels that become the *absolute* or *criterion* referents for judging the adequacy of skill related performance.

3. *Assessment.* Objectives define essential skills in terms of observable behaviors. They make possible the assessment of individuals' skill development status through the use of criterion referenced tests, i.e. tests devised explicitly to sample the behaviors described in the objective.

4. *Identification of Appropriate Teaching/Learning Activities.* Given behavioral descriptions of acceptable skill-related performance and knowledge of individuals' skill development status, a teacher is in a position to devise instruction appropriate to meet specific needs. In practice, this means that the teacher can select from the array of available instructional materials and activities those that appear to be most appropriate for a given student in a given situation at a given point in time.

5. *Evaluation.* Functional reading ability, the ability to cope with the reading tasks encountered in the real world is the end product. Evaluation amounts to constant checking to be assured that functional reading ability is attained. If not, then there is reason to examine all aspects of the approach in order to determine where the process has broken down.

The framework is summarized in the schema that follows. Note that a management element has been added to deal with some of the essential aspects of implementing the approach that is implicit in the framework.


MANAGEMENT
Individual records; Grouping; Scheduling; Organization of materials and procedures
ASSESSMENT
INSTRUCTION
SKILLS AND OBJECTIVES
EVALUATION

## COMPONENTS OF THE WISCONSIN DESIGN

In developing the Design, the skills included under a broad definition of reading were divided into six areas: Word Attack, Comprehension, Study Skills, Self-Directed Reading, Interpretive Reading, and Creative Reading. The skills were drawn from a carefully prepared curriculum guide and subsequently refined in view of feedback from teachers and reading specialists who have worked with the list in the field, extensive reviews of the related literature and instructional materials, the opinions of authoritative reviewers, and experience in pilot situations. The resulting list of skills serves as the basis for the entire Design.

Within the six areas the skills are clustered at levels that correspond generally to traditional grade levels. The six skill areas and traditional grade level equivalents of the skill levels are shown in Table 1. The grade designations are given only for transition purposes; they should be ignored as soon as a continuous progress approach to skill development becomes operational.

Table 1

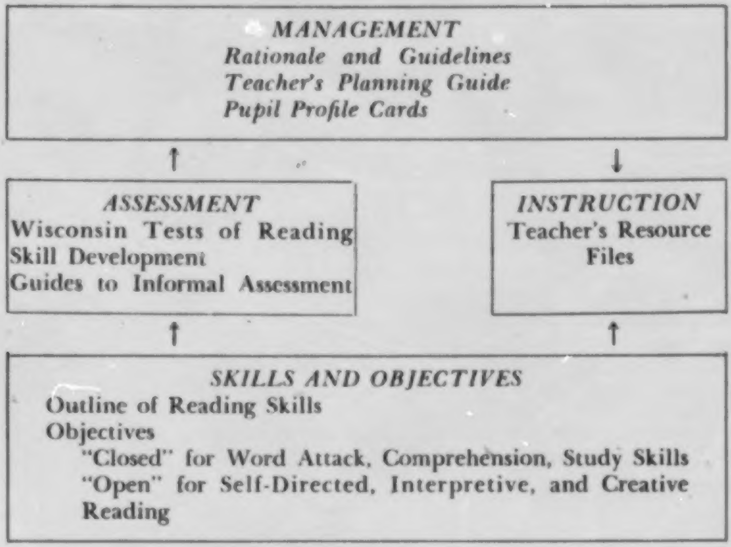
Skills by Area and by Traditional Grade Level

| | Grade | | | | | | |
|---|---|---|---|---|---|---|---|
| Skill Area | K | 1 | 2 | 3 | 4 | 5 | 6 |
| Word Attack | A | B | C | D | – | – | – |
| Comprehension | A | B | C | D | E | F | G |
| Study Skills | A | B | C | D | E | F | G |
| Self-Directed Reading | A | B | C | D | <——E——> | | |
| Interpretive Reading | A | B | C | D | <——E——> | | |
| Creative Reading | A | B | C | D | <——E——> | | |

In addition to the list of essential skills for each area, the Design includes a specific objective stated in behavioral terms for each skill in Word Attack, Comprehension and Study Skills areas and "open" objectives for the remaining areas; machine scorable criterion referenced tests and/or guides to informal assessment for each behavioral objective; pupil profile cards that permit the systematic grouping and re-grouping of pupils according to skill development needs; and resource files of suggested published materials and teaching procedures keyed to specific skills for use in instruction. Thus, the Design provides the basis for a skill centered approach to reading instruction, a means for monitoring individuals' progress in skill development, and a management system for both students and instructional materials. In addition, manuals are provided to assist with the implementation of the Design. A *Rationale and Guidelines* is addressed to personnel who will provide leadership in planning and providing for implementation of the Design. Underlying assumptions and requisites of a total readig program are included as well as specific guidelines for implementation. The *Teacher's Planning Guide* is limited to the specifics required for implementation in the classroom. Both manuals reflect several years' experience in schools that have collaborated in developing the Design. The procedures suggested are compatible with most but not tied to any particular instructional set-up.

In the schema that follows, components of the Design are placed in the framework. The assessment components permit teachers to focus upon behaviors related to specific skills, which in turn permits them to organize instruction around the specific skill development needs of individuals.

Components of the Design


*MANAGEMENT*
*Rationale and Guidelines*
*Teacher's Planning Guide*
*Pupil Profile Cards*
*ASSESSMENT*
Wisconsin Tests of Reading Skill Development
Guides to Informal Assessment
*INSTRUCTION*
Teacher's Resource Files
*SKILLS AND OBJECTIVES*
Outline of Reading Skills
Objectives
"Closed" for Word Attack, Comprehension, Study Skills
"Open" for Self-Directed, Interpretive, and Creative Reading

The Design amounts to a means for testing an important hypothesis derived from the basic assumption that underlies its development: A skill development focus in reading will result in improved overall reading achievement. The data that are available after one year of field testing of the word attack element are encouraging. Detailed reports of the field test results may be obtained from the Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin 53706.

EUNICE N. ASKOV
University of Wisconsin

# THE WORD ATTACK ELEMENT OF THE DESIGN[1]

The Word Attack element of the Wisconsin Design was the first to be completed. The formal field test of Word Attack began during the 1970-71 school year and was continued during 1971-72. Therefore, a discussion of the Word Attack components will reveal what also exists in the Study Skills element and what is planned for Comprehension.

## SKILLS AND OBJECTIVES

The foundation for the Word Attack element consists of the Outline of Reading Skills and related behavioral objectives. (The list of essential reading skills and the related objectives for Word Attack may be found in Otto and Askov, 1970.)

The list of Word Attack Skills was originally drawn from a carefully prepared curriculum guide of a public school system. The list has subsequently been refined in view of feedback from teachers and reading specialists who have worked with the list in the field, extensive reviews of the related literature and instructional materials, the opinions of authoritative reviewers, and experience in pilot situations. The skills, then, are the ones commonly taught in the elementary school and considered to be essential to reading development. The remaining components of the Design—the assessment exercises, the pupil accounting system, and the aids to instruction— are keyed to the specific Word Attack skills listed in the Outline.

---

[1] This is paper two (of five) from the symposium, A skill centered approach to reading instruction: The Wisconsin Design for Reading Skill Development.

Behavioral objectives were written for each Word Attack skill. The objectives were written at a mid-level of specificity, corresponding to Krathwohl's second level (1965). In other words, they fall between general program objectives at the most abstract level and specific objectives for instructional materials. The purpose of the behavioral objectives of the Design, then, is to define the tasks that the child must be able to perform in order to demonstrate skill mastery. Therefore, the objectives are specific enough to define the tasks but general enough to give the teacher considerable flexibility in designing an instructional program.

### ASSESSMENT

Each skill in Word Attack is assessed by a criterion-referenced test. In other words, performance on a test is judged by the extent to which the criterion behaviors are exhibited. The behavioral objectives for the skill define the criterion behaviors, and the criterion-referenced tst assesses the ability of the child to perform the tasks called for in the objectives. Since each child is evaluated only in terms of his ability to perform the criterion behaviors, no group-referent norms (e.g., percentile, stanine) exist. Ideally, a pupil would be expected to respond correctly to all of the items on any given test. In recognition of measurement limitations and situational variables, however, a criterion of 80% correct is frequently set.

Two types of assessment instruments exist. One type consists of informal guides to individual skill observation. They are intended to assist teachers in observing specific, skill-related behaviors in individuals and to serve as models for the development of additional exercises.

The odther type of assessment is the formal *Wisconsin Tests of Reading Skill Development*. A paper-and-pencil test exists for most skills in Word Attack. (For the few skills that are not amenable to written assessment individually administered performance inventories are provided.) Each test consists of approximately 15-20 items. Time limits are not imposed and examiners are encouraged to respond to questions regarding directions and unknown words in the tests. Each test has demonstrated reliability at a reasonably high level, usually having reliability coefficients of .80 or better.

### MANAGEMENT

To bridge the gap between assessment and instruction, a management system is provided. Some means of systematic record-keeping is essential if the results from assessment are to be applied in instruction. The Pupil Profile Card is the primary vehicle for pupil accounting with the Design. (See Figure 1 for an example of the card for Word Attack.)

All of the Word Attack skills are listed on the card. When there is evidence of mastery for a given skill, the hole at the edge of the

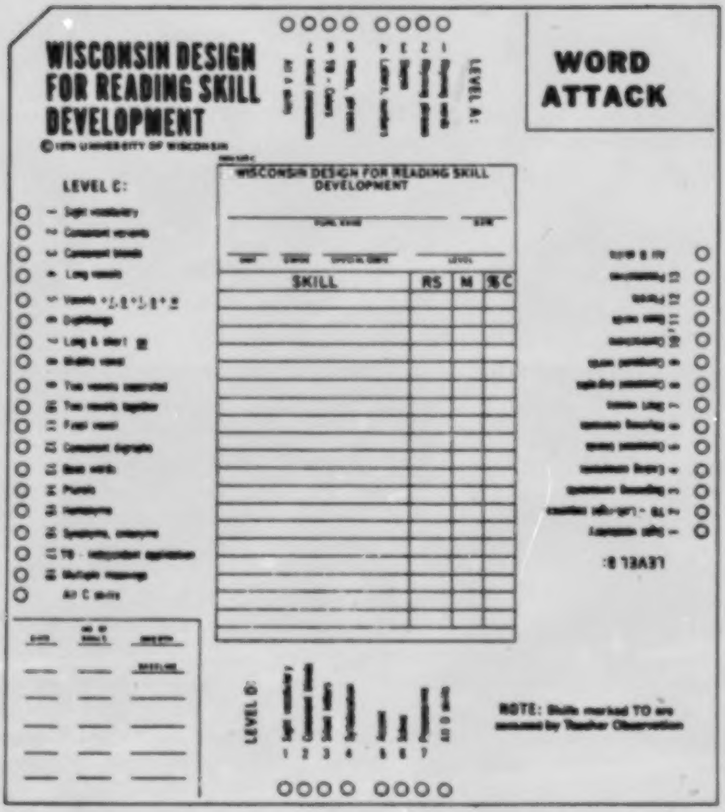

WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT
WORD ATTACK
LEVEL A:
LEVEL B:
LEVEL C:
LEVEL D:
WISCONSIN DESIGN FOR READING SKILL DEVELOPMENT
SKILL
RS
M
%C

Fig. 1. Individual Card for Card-Sorting Scheme

card next to the skill name is "notched." Thus, a pupil's skill development record amounts to "closed' 'holes for skills that need development and notched or "open" holes for skills that have been developed. To identify pupils with common skill development needs, the teacher selects a given skill, passes a skewer through the appropriate hole in a stack of cards, and shakes off the loose cards. The cards that remain on the skewer belong to pupils who need to work on the skill.

To assist with the mechanics of individual record keeping when test booklets are machine scored, a sumary of test results is provided as a part of the scoring service. The summary, which indicates what skills have been mastered, is printed on a gummed label that can be affixed to the center of a Pupil Profile Card.

## INSTRUCTION

Although plentiful and adequate instructional materials exist for teaching Word Attack skills, rarely are they systematically organized according to skills. Therefore, the *Teacher's Resource Files—Word Attack* list available instructional materials for each skill in the Outline of Reading Skills. A separate folder is thus provided for each skill. Page numbers in appropriate commercial materials for teaching a given skill are listed on the folder, and suggestions for teacher-directed activities and procedures are printed on sheets of paper which can be inserted inside the folder. Since not all material for teaching a skill could be included, the personnel in each school building are therefore encouraged to add to the files the kinds of materials and activities they find most useful.

## A FINAL WORD

Although undoubtedly some modifications in the Wisconsin Design materials would have to be made if they were used with adults, there is no reason to believe that the scope and sequence statement of essential skills for beginning adult readers would be different. Suggestions for use of the Wisconsin Design in Adult Basic Education programs may be found in Askov, Otto, and Fischbach (1970).

## REFERENCES

Askov, E. N., Otto, W., & Fischbach, T. Development of specific reading skills in adult education. In F. P. Greene (Ed.), *Reading: the right to participate*. National Reading Conference 20th Yearbook. Milwaukee: the Conference, 1971. Pp. 46-52.

Krathwohl, D. R. Stating objectives appropriately for program, for curriculum, and for instructional materials development. *Journal of Teacher Education*. 1965, *16*, 83-92.

Otto, W., & Askov, E. *The Wisconsin Design for Reading Skill Developments Rationale and Guidelines*. Minneapolis: National Computer Systems, 1970.

KARLYN KAMM
University of Wisconsin

# THE STUDY SKILLS COMPONENT OF THE DESIGN[1]

The focus of this paper is on the skills and objectives in the Study Skills area of the Wisconsin Design for Reading Skill Development. The intent is to describe the nature of these skills and the procedure used to identify them.

Our outline of skills and objectives is the result of an extensive search. The search included a survey of the professional literature, instructional materials—e.g., textbooks, kits, workbooks—and standardized tests. The information gained from this survey, together with the information gained from consultants and from consultation with school personnel is the basis for the list. As an organizational framework the study skills were divided into three subareas: maps, graphs and tables, and reference skills. The findings related to each of these subareas can be briefly summarized. A more detailed summary is contained in Otto, Kamm, and Weible (in press).

## MAPS

Though there has been an increased emphasis on map skills for the past 25 years, little attention has been paid to creating an organizational or developmental framework for the skills. Most of the skill outlines that have been devised suggest that there is a vast number of unrelated map skills, rather than that there are only a few fundamental skills, each with a number of related subskills. In addition to overlooking such a developmental framework, existing instructional materials tend to ignore completely important background concepts that a learner must understand before undertaking to learn a particular skill. Standardized tests seem to parallel the instructional materials in that they treat the skills as an array of unrelated tools the learner should acquire. The skills appear to be selected randomly for assessment. In general, our survey showed that up to the present time, little thought has been given to either the interrelationship existing among the map skills or to the related concepts. Teaching of the skills has been incidental rather than systematic.

---

[1] This is paper three (of five) from the symposium, A skill centered approach to reading instruction: The Wisconsin Design for Reading Development.

Our position is that a limited number of fundamental skills run through the subarea of maps at increasing levels of sophistication. In our outline these basic skills are introduced at the primary level and serve as building blocks on which more sophisticated variations of the same skills are added. We have made an attempt to deal not only with skills at increasingly more sophisticated levels, but also with the background concepts related to each skill (Askov, Kamm, Buschek, & Gustafson, 1971).

Three strands were identified in the subarea of maps: representation, location and measurement. A brief discussion of each strand and its divisions follows. A more detailed description is contained in the *Teacher's Planning Guide–Study Skills* (Askov and Otto, 1971).

Within the first strand, representation, two continua exist. The symbols themselves range along a continuum from concrete (e.g., pictures) to abstract (e.g., lines and dots). Likewise, the reality in the environment that is being represented ranges from concrete objects (e.g., a house) to area characteristics (e.g., an urban area). The developmental sequence, i.e., from concrete to abstract, is obvious in this strand.

Location comprises two divisions of skills, grid and direction. The progression in grid is from using a simple picture grid to using latitude and longitude to locate points. The skills in direction range from the use of general terms (e.g., up, over, beside) in the environment and in two-dimensional representations, to the use of cardinal and intermediate directions as applied to the globe, environment and maps.

Measurement also comprises two divisions, "the scale" and "to scale." Skills pertaining to "the scale" range from the use of crude approximation (e.g., long way) to precise computations. Skills pertaining to "to scale" involve the learner's understanding of the need for reduction in maps.

## GRAPHS AND TABLES

The critical question in teaching graphic materials is, What is the optimal sequence for teaching different kinds of graphs? Several viewpoints regarding the relative difficulty of picture, bar, circle and line graphs are presented in the literature. Perhaps the focal point, however, is not the sequencing of the graphs, themselves, but rather the sequencing of the skills involved in interpreting a particular graph. A natural and logical progression of skills begins with a simple, pictorially represented, one-to-one correspondence and moves to include a many-to-one correspondence as well as different ways to represent graphically the various correspondences. An example of one to one correspondence would be a graph which contains the same number of symbols as objects being represented. A many to one correspondence would show each symbol standing for more than one object.

These graph skills are closely tied with arithmetic skills, and in deciding upon a viable progression for an instructional group, care should be taken to see that there is a parallel between the mathematical background of the learners and the graph skills being introduced. For example, a circle graph, though generally thought to be quite sophisticated, could be introduced at the kindergarten level to show simple proportion i.e., a circle colored 25% green and 75% red, shows "many more red than green."

The content presented also plays a major role in deciding when a graph is appropriate for a given learner. The teacher, in addition to considering the skills necessary for interpretation, should also be sure the learner clearly understands the need for using the graph. Weintraub (1967) feels that, "Even though the research evidence may suggest one particular form as being simpler to interpret than another, the subject matter to be presented in graphic form must often dictate the most suitable graph for ease of interpretation [p. 349]."

In summary ,our survey points out that the developmental progression in graphs and tables should be dictated not by the type of graph or table, but rather by the particular skills necessary to the interpretation. Content seems to determine the appropriateness of the type of graph or table selected.

Our list of graph and table skills follows a developmental progression which begins with a simple, pictorially represented, one-to-one correspondence and includes a number of many-to-one correspondences. The four types of graphs—picture, bar, circle and line—and the single and multicolumn tables and schedules are introduced at various levels primarily to provide the child with practice on a variety of graphs and tables. The assumption is that appropriate context will be used for teaching each type of graph and table and the related skills.

### REFERENCE

In 1963 the National Council for the Social Studies published a revised and exhaustive document on study skills in its thirty-third year book, *Skill Development in Social Studies*. The book contains a well organized list of 180 study skills and points out the complexity of the study skills as a whole and provides an overview of the field.

Because this great number of skills can be distinguished and because most of these skills cross subject areas, the search in reference skills included many state curriculum guides and well-developed lists of objectives for particular school systems—such as Oak Park, Illinois—as well as textbooks and instructional materials.

There are many differences between the skills as they are presented by various authors. Grade level placement, terminology (terms such as "introduce," "main emphasis," and "planned readiness" are unacceptable in terms of meaningful behavioral objectives), assigning responsibility for both teaching and evaluating mastery of reference skills, and the spiral pattern of teaching every skill each year with no guides as to what can be mastered at various levels of the spiral are four examples of these differences.

Thus, in the area of reference skills, it is easy to find literature which describes the importance of the study skills, but it is more difficult to find sources which define the skills adequately and it was nearly impossible to arrive at a consensus as to skill placement. Our approach, then, was first to identify essential skills and then to trace the spiral of development in order to find various check points along the skill development sequence. These check points eventually became either the teacher observations or the paper and pencil tests constructed for the Design. 'Correctness' of placement has been checked with some textbooks, but there is no uniform agreement as to grade placement in textbook series, conferences with teachers, and the eventual process of pupil try-out itself. From the array of skills listed and described in the various sources, we selected those that were either mentioned more frequently or those that fell within certain frequently mentioned organizational patterns. Our next step was to create an organizational structure around each skill we considered to be essential. This was done by following through the steps of searching for information, considering whether the information met the need, and, finally, using that information.

The reference subarea is divided into seven strands. Each strand has a variety of skills within it, and each follows a spiral pattern of using more sophisticated skills with increasing independence at each higher level. "Location" involves alphabetizing skills and their application in the library. "Book skills" range from holding a book properly to knowing the idiosypcratic organization of a variety of references. In "Library use" the learner progresses from learning to checking out books to using the Dewey Decimal System. "Reference materials" includes skills ranging from a picture dictionary to reference indexes. The assumption in "Reading rate and purpose" is that once decoding words has ceased to be a major focus of the reading task, it is appropriate to begin to develop rate flexibility. "Organizing and evaluating materials" include basic sequencing skills and drawing conclusions and checking accuracy and relevancy of materials. "Work Habits" range from following simple directions to engaging in independent research.

### SUMMARY

There are two assumptions which underlie our fundamental skills and subskills framework in the three subareas: that learning can best be achieved through the study of interrelated skills rather than isolated skills; and having the skills presented in a developmental sequence facilitates their being taught in context, e.g., map skills are an integral part of the social studies program. Creating the need to learn a skill is essential if learning and understanding is to be complete.

The list of skills and objectives for the study skills area of the Wisconsin Design for Reading Skill Development can be found in the *Teacher's Planning Guide–Study Skills.* Other components of the Design for the Study Skills area include the *Teacher's Resource Files–Study Skills,* criterion-referenced machine scorable tests, and the Pupil Profile Cards. The purpose of these components can be inferred from the previous discusion of Word Attack.

### REFERENCES

Askov, E., Kamm, K., Buschek, P., & Gustafson, D. *Teacher's Resource Files—Study Skills.* Minneapolis: National Computer Systems, 1971.

Askov, E., & Otto, W. *The Teacher's Planning Guide—Study Skills.* Minneapolis: National Computer Systems, 1971.

Otto, W., Kamm, K., & Weible, E. *Wisconsin Design for Reading Skill Development: Rationale and Objectives for the Study Skills Element.* Madison: Wisconsin Research and Development Center for Cognitive Learning, in press.

Weintraub, S. What research says on the reading teacher. *Reading Teacher,* 1967, *20,* 345-349.

ROBERT D. CHESTER
University of Wisconsin

# THE COMPREHENSION COMPONENT OF THE DESIGN[1]

The comprehension component of the Wisconsin Design for Reading Skill Development is the last of the three major components to be developed. The rationale and objectives for this part of the Design are due to be completed by the end of January with test construction beginning around February 1. Although it is premature at this time to attempt to outline specific skill objectives, it is possible to list skills for which objectives are being developed and to say something about the rationale for choosing them. That, then, is the purpose of this paper.

Thorndike's "Reading as Reasoning" (1917) is probably the most important of the early studies in comprehension. The essence of Thorndike's theory is provided in the following quote taken from his study.

> Understanding a paragraph is like solving a problem in mathematics. It consists in selecting the right elements of the situation and putting them together in the right relations, and also with the right amount of weight or influence or force for each. The mind is assailed as it were by every word in the paragraph. It must select, repress, soften, emphasize, correlate and organize, all under the influence of the right mental set or purpose or demand (p. 329).

---

[1] This is paper four (of five) from the symposium, A skill centered approach to reading instruction: The Wisconsin Design for Reading Development.

Although one might broadly interpret Thorndike's study as an attempt to examine the components of comprehension, it was not until sometime later that significant attempts were made toward isolating these components.

Some of the more important of these early attempts were carried out by Richards (1929), Berry (1931), Touton and Berry (1931), Dewey (1935), and Traxler (1941). Most of the studies were based on correlational type studies or logical guesses. The list of subskills produced was extensive and quite proably has led to some of the confusion in terminology which still exists today. Lists ranged from Berry's list of sixteen skill characteristics to Dewey's list of two.

It was not until the late 30's or early 40's that the procedures of factor analysis were applied to the problems of isolating the various subskills of comprehension. Although Davis (1944) is generally credited with the most important of the early factor analysis of comprehension skills, it is worth noting that at least three significant factor analysis studies had preceded his. As early as 1938, Feder, using the factor analysis techniques analyzed a matrix that included tests of factual reading, inference reading, appreciation, and speed of reading. From specially built comprehension maturity tests, Feder obtained data indicating that reading for information and reading for inference are relatively independent skills.

A second factor analysis study worth noting was reported by Langsam (1941). Using college freshmen as subjects, she administered reading tests yielding fourteen scores and an intelligence test that yielded seven scores. Applying Thurstone's Controid Method of factor analysis she was able to identify five factors: verbal; perceptual; word; number; and seeing relationships.

A third factor analysis study which was carried out but not published was that of Pankaskie (1940). This analysis of the Iowa Silent Reading Test, Iowa English Training Examination and Iowa Mathematics Aptitude Test showed three factors, i.e., speed of comprehension, vocabulary, and ability to find answers to questions. Subtests of the Iowa Silent Reading Tests were not found to be 'pure' measure of these factors.

Undoubtedly the best known of the early factor analytic studies of comprehension was that of Davis (1944). His effort was the first that made use of "tests especially constructed to measure the mental skills in reading regarded as of greatest importance by authorities in the field [p. 185]." After an extensive review of the literature in the field, he came up with a list of nine testable skills:

1. Knowledge of word meanings
2. Ability to select the appropriate meaning for a word or phrase in the light of its particular contextual setting
3. Ability to follow the organization of a passage and to identify antecedents and references in it

4. Ability to select the main thought of a passage
5. Ability to answer questions that are specifically answered in a passage
6. Aiblity to answer questions that are answered in a passage but not in the words in which the question is asked
7. Ability to draw inferences from a passage about its contents
8. Ability to recognize the literary devices used in a passage and to determine its tone and mood
9. Ability to determine a writer's purpose, intent, and point of view, i.e., to draw inferences about a writer

Davis computed the intercorrelation among the nine tests. He interpreted a factor analysis of the results as indicating the presence of nine factors. Six of these nine were subsequently established as being significant at the .05 level. However, of the nine factors, word knowledge accounted for the greatest part of the variance. Reasoning in reading ranked second followed by the literal meaning factor. Even though the factor analysis did identify six components, Davis concluded that only components I and II (word knowledge and reasoning in reading) "are measured with sufficient reliability to warrant their use for practical purposes [p. 194]."

Thurstone (1946) reanalyzed Davis' data by Spearman's Unidimensional method. He concluded that a single common factor (reading ability) accounts for the correlations among the tests with residuals remarkably small in view of the fact that the tests were designed to test nine supposedly different skills.

Harris (1948) in a study equivalent to that of Davis in nature and scope arrived at conclusions similar to those of Thurstone.

> From this study it is concluded that one general factor is adequate to account for the intercorrelations of the seven variables, each of which specifies a particular behavior or skill that was defined as an aspect of comprehension of litreature. It should be noted particularly that group factors were not found (p. 339).

In 1952 Hunt carried out a study similar to that of Davis. He re-examined the six skills earlier identified by Davis as being fundamental to six factors of reading comprehension. He determined that for the comprehension skills considered, other than vocabularly, the average differences between the point biserial values for each pair of skills measured are probably chance differences. He concludes, "It would appar, then, that what is being measured by each of the six skills examined in the study is to a large extent general reading comprehension and to a lesser extent, vocabularly or isolated word knowledge [p. 127]."

Hall and Roinson (1945), on the other hand, found quite different results when they factor analyzed the data from a battery of tests that included twenty-five separate reading measures, many of which were tests of reading non-prose material. The findings of their study showed the existence of six factors, one of which was not identifiale. These were:

1. Attitude of Comprehension Accuracy
2. Rate of Inductive Reading
3. Verbal or Word Meaning
4. Not Identifiable
5. Rate for Unrelated Facts
6. Chart-reading Skill.

Further justification for the multi-factor theory of comprehension is supplied by Vernon (1962). He found general comprehension and vocabulary to be separate factors. He also notes, on the basis of his data, that a distinction can be made between inferential and factual comprehension.

Probably the strongest support for a multi-factor theory of comprehension may be seen in Davis' (1958) most recent investigation. Using a set of carefully constructed tests designed to measure specific aspects of reading comprehension, Davis reinvestigated the unique variances of subtests of reading comprehension. However, in this study he used "a cross-validated unique analysis based on large samples of examiners [p. 510]."

Davis concluded, on the basis of a rather impressive study, that "comprehension among mature readers is not a unitary mental skill or operation [p. 510]." His factor analysis indicated that 32%, the largest percentages of unique variance, occurred in the case of Skill 1 (memory for word meaning). The second largest percentage of unique variance, 20%, occurred in the case of Skill 5 (drawing inferences from content). Three other skills which showed appreciable percentages of unique variance were Skill 8 (following the structure of the passage); Skill 6 (recognizing a writer's purpose, attitude, tone, and mood); and Skill 3 (finding answers to questions asked explicitly or in paraphrase).

Although the studies above do not exhaust the literature dealing with reading comprehension, they do, nevertheless, constitute a representative sample of the research in this area. Even from this limited review of the literature, the lack of consistency in findings regarding generalities and specificities of comprehension abilities is obvious. There is widespread disagreement among educators, not only as to the specific subskills of comprehension, but also as to whether indeed such subskills exist at all.

However, despite the confusion, there does appear to be emerging a growing amount of evidence in support of a multi-skill theory of comprehension. The majority of the research indicates that most of the measurable variance in tests of reading comprehension can be accounted for in terms of a small number of factors, perhaps six to eight. An examination of the research literature indicates that most of the so called comprehension sub-skills are not discrete skills but rather, are superficially varied interpretations of the same skill. It is this multi-skill theory of comprehension that we have chosen to accept, with the full knowledge that not everyone will agree.

One further decision of importance should also be noted, and that is the decision to use an operational definition of comprehension. The most obvious reason is that it defines and limits the area with which we will be attempting to deal. A second and equally obvious reason is that it would be virtually impossible to develop tests and materials that are definitive according to everyones specifications. Unfortunately, this has not stopped a number of people in the past and this leads to our third reason for operational definitions. We hope to lessen, or at least, not add to the confusion already so prevalent in the area of comprehension.

In devleoping the comprehension component of the Design, three areas have been selected for focal interest because of thir historical, empirical, and logical importance. Their dimensions encompass most of what has heretofore been called comprehension, yet they can be examined and developed more readily than their multitudinous components. Seen as three factors, they are a verbal factor, a literal factor, and an inferential factor.

For purposes of development, the verbal factor is submerged in the other two factors. A discussion of the rationale for this decision can be found in a forthcoming position paper (Chester & Kamm, 1972). For the time being, it seems sufficient to say simply that the verbal factor is not seen as a skill, but rather as accumulated knowledge. Also, because of obvious space limitations, there will be no attempt to state here the rationale for each separate subskill of comprehension. This too may be found in the position paper.

For purposes of this paper, we simply note that the comprehension element for the Design has been developed under two types of skills, i.e., literal skills and inferential skills. These specific skill areas are:

| *Literal* | *Inferential* |
|---|---|
| 1. Syntactical Structure | 1. Main Idea |
| 2. Sequence | *2. Meaning from Context |
| 3. Punctuation | 3. Evaluation (relevance) |
| *4. Affixes (semantics of) | 4. Conclusions (reasoning) |

* testing skills involving verbal factor

A legitimate question might be, how is the Design approach different from other approaches already in existence. The answer is obvious; it isn't in many ways. It tests many of the traditional skills obviously present on several other existing tests. However, it is different in a number of real ways.

For one thing, the Design approach is based on diagnostic tests. This in itself implies that it will deal only with skills which can be identified, measured, and taught. For instance, even though vocabulary knowledge can be identified, and to an extent measured, it is not a skill and cannot be taught as such. Only those vocabulary skills involved which enable a student to gather meaning from a word can be taught. This is also the case with other areas such as appreciation, value judgment, pure inference, etc.

A second difference lies in the organization of the skills in the Design. We are all aware of the fact that names are misleading—especially where tests are involved. The skill which one test labels general comprehension may be called reading to obtain facts by another test and referred to as testing meaning from context by still another. The Design, through operational definitions and objective centered criteria, is attempting to avoid the game of misnomers and to subsume the actual skills involved under a few common demoninators. Thus what has been traditionally called reading for facts, ideas, supporting details, main ideas, details, organization, sentence meaning, sequential order, etc., is all subsumed by syntactical structure, sequence, and main idea. Similarly, such tests as anticipating outcomes, predicting conclusions, inference, translating ideas and making inference, cause and effect, reasoning, interpreting, anticipating, etc., are all subsumed under drawing conclusions.

In conclusion, while recognizing that there are a number of conflicting points of view concerning comprehension, the Design is attempting to capitalize on commonalties rather than diversities. Through operational definitions, objective centered criteria, and skills selection based on historical, empirical, and logical support, we are attempting to identify and measure comprehension skills which can be taught, and to facilitate the teaching of these skills.

## REFERENCES

Berry, B. T. Improving freshman reading habits. *English Journal*, 1931, *20*, 824-828.

Chester, R. D., & Kamm, K. Wisconsin Design for Reading Skill Development: rationale and objectives ofr comprehension element. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

Davis, F. B. Fundamental factors of comprehension in reading. *Psychometrika*, 1944, *9*, 185-197.

Dewey, J. C. The acquisition of facts as a measure of reading comprehension. *Elementary School Journal*, 1935, *35*, 346-348.

Feder, D. D. Comprehension maturity tests in a new technique in mental measurement. *Journal of Educational Psychology*, 1938, *29*, 597-606.

Hall, W. E., & Robinson, F. P. An analytical approach to the study of reading skills. *Journal of Educational Psychology*, 1945, *36*, 429-442.

Harris, C. W. Measurement of comprehension of literature: II Studies of measures of comprehension. *School Review*, 1948, *56*, 332-342.

Hunt, L. C. A further study of certain factors associated with reading comprehension. Unpublished doctoral dissertation, Syracuse University, 1952.

Langsam, R. S. A factor analysis of reading ability. *Journal of Experimental Education*, 1941, *10*, 57-63.
Pankaskie, M. Some factors in reading achievement at the college level. *Psychological Bulletin*, 1940, *37*, 473. Abstract.
Richards, I. A. *Practical criticism*. New York: Harcourt Brace, 1929.
Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, 1917, *8*, 823-832.
Thurstone, L. L. Note on a reanalysis of Davis' reading tests. *Psychometrika*, 1946, *11*, 185-188.
Touton, F. C., & Berry, B. T. Reading comprehension at the junior college level. *California Quarterly of Secondary Education*, 1931, *6*, 245-251.
Traxler, A. E. A study of the Van Wagenen-Dvorak Diagnostic Examination of Silent Reading Abilities. Educational Records Bulletin No. 31. New York: Educational Records Bureau, January, 1941, 33-41.
Vernon, P. E. The determinants of reading comprehension. *Educational and Psychological Measurement*, 1962, *22*, 269-286.

ROBERT T. RUDE[1]
University of Wisconsin

# IMPLEMENTATION AND FIELD TESTING OF THE DESIGN[2]

The formal field test of the Word Attack began during the 1970-71 school year and is continuing during the 1971-72 school year. The Study Skills and Comprehension components will be field tested at a later time and will not be reported in this paper.

The Word Attack component of the Design is presently being field tested at the primary levels in eighteen Wisconsin and five Colorado elementary schools. Five of the eighteen Wisconsin schools are located in the central-city portion of Milwaukee. These sites are referred to as disadvantaged since median reading achievement is a year or more below grade level. The remaining Wisconsin and Colorado sites are referred to as advantaged schools. This paper will report on the performance of pupils in these schools who had experienced the program for approximately six months.

## OBJECTIVES FOR THE WORD ATTACK COMPONENT

Three major objectives have been stated for the Word Attack component:

### *Terminal Objective*

> The student upon attainment of all Level D Skills will be able to attack independently, phonically and/or structurally

---

[1] I am indebted to Mrs. Mary Quilling for her technical assistance in preparing this paper.

[2] This is paper five (of five) from the symposium, A skill centered approach to reading instruction: The Wisconsin Design for Reading Development.

regular words and will recognize on sight all the words on the Dolch list. Children of average or above average ability will attain this objective at least by the end of the fifth year (fourth grade) in school, while others will attain this objective by the end of the seventh year.

*Interim Objective*

The student upon attainment of all Level C Skills will be words and their structural variants, and will recognize on sight able to attack independently, phonically regular one syllable the easier 110 words on the Dolch list. Children of average or aobve average ability will attain this objective by the end of the fourth year in school, while others will attain this objective by the end of the fifth year in school.

*Specific Objectives*

Forty-five specific behavioral objectives have been stated elsewhere (Otto and Askov, 1970). The student whose pupil record indicates he has attained a particular skill will perform at a mastery level (80% or higher) on a subsequently administered criterion-referenced test.

### OBJECTIVES OF THE EVALUATION

Three additional objectives have been stated for the evaluation of the overall program. These objectives are:

(a) to determine whether the terminal, interim, and specific objectives of the program are met. (Attainment of terminal and interim objectives is meaningfully assessed only in the second year of program operation and thereafter.)

(b) to determine whether the Word Attack and other reading skills of children using the program are greater than those of children not using the program in the same schools and at the same age-grade levels in the year prior.

(c) to learn how the program is implemented in a variety of selected sites in which leaders are not exposed to a standard training program and how local variants affect program success.

Preliminary data related to the attainment of the Word Attack terminal and interim objectives are presently being gathered and therefore will not be reported in this paper. Subjects in the field test are presently engaged in mastering the forty-five specific behavioral objectives.

At this point in time, it is perhaps more crucial to examine the objectives of the evaluation. As specified earlier, attainment of the program objectives is assessed meaningfully only in the second year of program operation and thereafter. The latter two evaluation objectives, performance of students who were involved in the program versus those who were not and program implementation practices, are the primary foci of this paper.

## DESIGN OF THE STUDY

During May, 1970, randomly selected subjects in the twenty-three field test schools who had not experienced the program were administered selected batteries of the *Wisconsin Tests of Reading Skill Development—Word Attack* as well as three standardized reading tests. The standardized reading tests were the *Cooperative Primary Tests,* Word Analysis (1965); the *Stanford Achievement Test,* PI and PII (Kelly, Madden, Gardner, and Rudman, 1964); and the *Comprehensive Tests of Basic Skills,* Vocabulary and Comprehension (1968). To facilitate the testing, students were randomly assigned to sittings and ordinarily participated in only one test sitting. The tests which comprise each sitting are found in Appendix A. During May, 1971, another random sample of students was selected and participated in the interim testing program.

While the unit of reference is the individual student, the unit for statistical analysis must be the entire grade cohort. This is necessary since the presence or absence of the program is experienced for all or none of the cohort and the effects of the particular conditions of implementation are experienced by the entire participating school staff.

To determine whether a change in performance betwen the control group (subjects not exposed to the program) and the experimental group (subjects exposed to the program) can be inferred, the data were used to test the null hypothesis that no change had occurred. Hypotheses concerning change at each grade level, and within grade level, by type of measure, were tested separately. In other words, for a given grade level, up to three hypotheses related to no change on program-dependent Design subtests, program-independent Word Attack tests, and standardized measures of general reading ability were tested. In each case, the data consisted of one or more change scores for each school in the field test. Each change score represented the difference between the mean scores for the first implementation year and the preimplementation year for a given measure or set of measures. (Composite scores were formed for the three to five Design subtests administered in a single setting.)

In those cases for which there is but one change score per school for a particular grade and content type, a two-tailed t-test was used to determine whether the hypothesis of no change should be rejected or not rejected. If the hypothesis is rejected, and if the change is positive, a gain was inferred; if significantly negative, a loss was inferred.

A more complex analysis was performed when two or more change scores were available from each school. The three-step analysis first require rescaling of the change scores to take into account scale difference between measures (or their composites) and to give more weight to the more reliable measures. Each raw change score was multiplied by a constant inversely proportional to the standard deviation of the schools' change scores.

Next, it was determined whether the amount and direction of change varied by measure, as rescaled. The expectation was that change would be consistent for the various measures of a given kind. To test the hypothesis, contrasts between changes in the various measures were firmed, and a t-ratio or Hotelling's $T^2$ statistic computd. Rejection of the hypothesis indicated that degree and direction of change was systematically different for the instruments, whereas from failure to reject supported the assumption of consistency in change.

Depending upon the outcome of the test of the consistency hypothesis, the last step varied. If consistency was inferred from step two in the analysis, then a $T^2$ test is applied to the average change score across measures to determine whether the raw unit change is nonzero. Descriptive statistics indicate the direction and amount of change. If change is not consistent across measures, the change scores ofr all measures are analyzed by Hotelling's $T^2$ test to determine whether a change occurred. If this test yields significant results, univariate tests are subsequently used to identify the measures in which there was change.

Program implementation practices, the third evaluation objective, were ascertained by randomly selecting a sample of teachers and mailing them questionnaires.

## FINDINGS

The most marked effects of the program was related to tests associated with the specific program objectives. Achievement of subjects enrolled in the program compared to those not enrolled favored the former group on more than ninety percent of the objectives. This finding pertained to both advantaged and disadvantaged schools.

In addition, performance on the standardized reading test was as good or better for participating pupils. Mean gains of six, nine, and thirteen raw score points in some inner-city schools on the *Cooperative Primary Test*—Phonics Analysis were observed. Increases for other tests were more modest.

Performance on vocabulary and comprehension measures of reading achievement were slightly negative for the former and negligible for the latter. These results may be attributable to a conversion from a general word recognition teaching technique to a phonic approach and emphasis on word attack skills rather than on comprehension. Increased scores on both vocabulary and comprehension are anticipated as the field testing continues.

Analysis of returned teacher questions revealed that about sixty percent of the teachers are responsible for instructing a single skill group, thirty percent for two groups, and ten percent for three groups. Approximately nine percent of the children in a skill group

Table 1

Summary of Significant Results for First Year Word Attack Field Test

| | Disadvantaged | | Advantaged | |
|---|---|---|---|---|
| Grade | WTRSD | Stand. | WTRSD | Stand. |
| 1 | $<.05$ | None | $<.01$ | $<.05^{d}$ |
| 2 | $<.01$ | $<.05^{a}$ | $<.01$ | N.S.$^{e}$ |
| 3 | N.S. | N.S.$^{b}$ | $<.01$ | $<.01$ (neg.)$^{c}$ |
| 4 | N.S. | N.S.$^{c}$ | None | None |

[a] Cooperative Primary Tests—Word Analysis
[b] Cooperative Primary Tests—Word Analysis
Stanford Achievement Tests—Word Meaning and Word Study Skills
[c] Comprehensive Tests of Basic Skills—Vocabulary and Comprehension
[d] Cooperative Primary Tests—Word Analysis
Stanford Achievement Tests, PI—Word Reading and Word Study Skills
[e] Stanford Achievement Tests, PII—Word Study Skills and Word Meaning

were excused before the last day because they had demonstrated skill mastery. Interestingly, twenty percent of the children in a skill group receive different instruction even though all children in a skill group demonstrate a common need. This indicates that individualization of instruction is occurring within the skill group.

Most students demonstrate skill mastery by performance on a paper-and-pencil test; only ten percent of the students are assessed strictly by teacher judgment. The majority of teachers administer tests themselves while two-thirds of the teachers perform their own scoring of tests.

The first-year field test of the Word Attack component has been completed while the second-year of the Word Attack and first-year of the Study Skills components has only begun. The Design, then, is in the first phase of an extensive field testing program. Additional field test results will be reported at professional meetings as well as through the Research and Developmest Center for Cognitive Learning publications (Quilling and Otto, 1971, in press).

## REFERENCES

*Comprehensive Tests of Basic Skills.* Monterey: California Test Bureau, 1968.

*Cooperative Primary Tests—Reading.* Princeton: Educational Testing Service, 1965.

Kelly, T., Madden, R., Gardner, E., & Rudman, H. *Stanford Achievement Tests*, PI and PII. New York: Harcourt, Brace & World, 1964.

Otto, W., & Askov, E. *The Wisconsin Design for Reading Skill Developments Rationale and Guidelines.* Minneapolis: National Computer Systems, 1970.

Otto, W., Kamm, K., Peterson, J., Harris, M., & Miles, P. *Wisconsin Tests of Reading Skill Development—Word Attack.* Minneapolis: National Computer Systems, 1970.

Quililng, M., & Otto, W. *The Large Scale Field Test of the Wisconsin Design for Reading Skill Development: An Interim Report.* Madison: Wisconsin Research and Development Center for Cognitive Learning, 1971.

## Appendix A

### Part 1

### Schedule of Sittings Where Present Reading Achievement is at or Above Grade Level

| Grade | Sitting | Test | Administration Time (In Minutes) |
|---|---|---|---|
| 1 | 1 | Cooperative Primary Tests<br>Word Analysis | 40 |
| 1 | 2 | Stanford Achievement PI<br>Word Reading<br>Word Study Skills | <br>15<br>25 |
| 1 | 3 | Wisconsin Tests of Reading Skill Development—Level B<br>Test 3—Beginning Consonant Sounds<br>Test 5—Consonant Blends<br>Test 6—Rhyming Elements | <br>10<br>10<br>10 |
| 1 | 4 | Wisconsin Tests of Reading Skill Development—Level B<br>Test 7—Short Vowels<br>Test 10—Contractions<br>Test 11—Base Words and Endings | <br>10<br>10<br>10 |
| 2 | 1 | Stanford Achievement PII<br>Word Study Skills<br>Word Meaning | <br>30<br>12 |
| 2 | 2 | Wisconsin Tests of Reading Skill Development—Level C<br>Test 3—Consonant Blends<br>Test 4—Long Vowel Sounds<br>Test 12—Consonant Digraphs | <br>10<br>10<br>10 |
| 2 | 3 | Wisconsin Tests of Reading Skill Development—Level C<br>Test 5—Vowel + r, a + l, a + w<br>Test 6—Diphthongs<br>Test 16—Synonyms and Antonyms | <br>10<br>10<br>10 |
| 2 | 4 | Wisconsin Tests of Reading Skill Development—Level D<br>Test 2—Three Letter Consonant Blends | <br>10 |
| 3 | 1 | Comprehensive Tests of Basic Skills<br>Vocabulary<br>Comprehension | <br>15<br>34 |
| 3 | 2 | Wisconsin Tests of Reading Skill Development—Level D<br>Test 2—Three Letter Consonant Blends<br>Test 3—Silent Letters<br>Test 4—Syllabication<br>Test 5—Accent<br>Test 7—Possessives | <br>10<br>10<br>10<br>10<br>10 |

## Part 2

## Schedule of Sittings Where Present Reading Achievement Is Below Grade Level, 1970

| Grade | Sitting | Test | Administration Time (In Minutes) |
|---|---|---|---|
| 1 | 1 | Wisconsin Tests of Reading Skill Development—Level A | |
| | | Test 1—Rhyming Words | 20 |
| | | Test 5—Words and Phrases | 15 |
| | | Test 7—Initial Consonants | 15 |
| 1 | 2 | Wisconsin Tests of Reading Skill Development—Level B | |
| | | Test 3—Beginning Consonant Sounds | 10 |
| | | Test 5—Consonant Blends | 10 |
| | | Test 6—Rhyming Elements | 10 |
| 1 | 3 | Wisconsin Tests of Reading Skill Development—Level B | |
| | | Test 8—Consonant Digraphs—sh, ch, th | 10 |
| | | Test 9—Compound Words | 10 |
| | | Test 12—Plurals | 5 |
| 2 | 1 | Cooperative Primary Tests | |
| | | Word Analysis | 40 |
| 2 | 2 | Wisconsin Tests of Reading Skill Development—Level B | |
| | | Test 3—Beginning Consonant Sounds | 10 |
| | | Test 5—Consonant Blends | 10 |
| | | Test 6—Rhyming Elements | 10 |
| 2 | 3 | Wisconsin Tests of Reading Skill Development—Level B | |
| | | Test 7—Short Vowels | 10 |
| | | Test 10—Contractions | 10 |
| | | Test 11—Base Words and Endings | 10 |
| 2 | 4 | Wisconsin Tests of Reading Skill Development—Level C | |
| | | Test 3—Consonant Blends | 10 |
| | | Test 4—Long Vowel Sounds | 10 |
| | | Test 12—Consonant Digraphs | 10 |
| 3 | 1 | Cooperative Primary Tests | |
| | | Word Analysis | 40 |
| 3 | 2 | Stanford Achievement Tests | |
| | | Word Meaning | 12 |
| | | Word Study Skills | 30 |
| 3 | 3 | Wisconsin Tests of Reading Skill Development—Level C | |
| | | Test 3—Consonant Blends | 10 |
| | | Test 4—Long Vowel Sounds | 10 |
| | | Test 12—Consonant Digraphs | 10 |

(Part 2, Continued Next Page)

(Part 2 Continued)

| Grade | Sitting | Test | Administration Time (In Minutes) |
|---|---|---|---|
| 3 | 4 | Wisconsin Tests of Reading Skill Development—Level C | |
| | | Test 5—Vowel + r, a + l, a + w | 10 |
| | | Test 6—Diphthongs | 10 |
| | | (D) Test 2—Three Letter Consonant Blends | 10 |
| 4 | 1 | Comprehensive Tests of Basic Skills | |
| | | Vocabulary | 15 |
| | | Comprehension | 34 |
| 4 | 2 | Wisconsin Tests of Reading Skill Development—Level C | |
| | | Test 5—Vowel + r, a + l, a + w | 10 |
| | | Test 6—Diphthongs | 10 |
| | | Test 16—Synonyms and antonyms | 10 |
| | | (D) Test 12—Three letter Consonant Blends | 10 |
| 4 | 3 | Wisconsin Tests of Reading Skill Development—Level D | |
| | | Test 3—Silent Letters | 10 |
| | | Test 4—Syllabication | 10 |
| | | Test 5—Accent | 10 |
| | | Test 7—Possessives | 10 |

EDWARD FRY
Rutgers University

# TRAINING READING TEACHERS USING DIRECT EXPERIENCE READING TASKS[1]

This is the first report of a curriculum development project aimed at improving the teaching of reading methods courses for teachers. Besides this section, there are four other sections, each written by those who worked on the project.

The basic premise is that the traditional lecture, textbook, and term paper, need to be supplemented by live experiences or manipulation of data. For example, we felt that it was better for a student to work a readability formula rather than just hear a lecture on how they work. We felt that an observation of eye movements is a better learning experience than reading about them in a textbook. The same is true for a host of other tasks, such as cloze technique, generative grammar, and oral reading measurement.

Most of these tasks have been tried out on graduate and undergraduate courses in reading, and they seem to be equally valuable at both levels. Perhaps, however, if they were more universally used in undergraduate courses, they would be less necessary at the graduate level. Sometimes the distinction between graduate and undergraduate level courses becomes fuzzy in the changing field of education.

In any event, even though some of the tasks require considerable work, we did not receive any loud complaints from students, and we hope that this is because the students really found that they were learning something.

---

[1] This is paper one (of five) from the symposium, Training reading teachers without lecture or textbooks.

My specific assignment in this presentation is to mention the Readability and Frequency Count Tasks. Other areas will be covered by the next four writers.

## DIRECT EXPERIENCE READING TASKS: READABILITY AND FREQUENCY COUNT

Readability formulas have had a continuing interest for educational researchers since the 1940's, but despite the longevity of this interest, there are still many teachers, perhaps the majority, who have, at best, only a cursory acquaintance with them. It is true that some of the early formulas had mildly imposing arithmetic calculations, but underneath that, most formulas are based on a measure of grammatical complexity and a measure of vocabulary difficulty.

### *The Readability Task*

My Readability Graph is used as a task because it has simplified these two important variables to a count of sentences and a count of syllables per hundred words. The mathematical calculations are mostly solved by using a graph to compute grade level. The student is asked to determine the readability of a sample passage (with the correct answer given in the key); then he is asked to work a readability formula on two children's textbooks of his choice.

References are given to the Spache and Dale Chall formulas for students who wish to look at or work those two well-known measures of readability, or instructors who wish to assign them.

### *Frequency Count Tasks*

The frequency count tasks are probably among the most time-consuming in the series of tasks. The student is asked to do a frequency count of 1500 running words which is enough to show him that certain words do have a much higher frequency than others. Hopefully, this observation and a careful look at the phonetic regularity of the highest frequency words will lead him to some appreciation for the 'word' or 'sight' method of teaching reading.

Later, a frequency count is made of phoneme-grapheme correspondences which should show something of the relative importance or unimportance of many phonic generalizations. Not only must the student teacher learn a good bit about phonics in order to do the count, he might also be faced with something like the problem of the pupil who is glibly told to "Sound out the words you don't know."

With this as an introduction to the Direct Experience Reading Tasks, we hope you will see that the basic teaching premise is rather like that of the laboratory section of a chemistry course; namely, to present the student with a bit of reality which might have a bit more punch than a verbal discussion by the instructor. The Tasks are not yet refined, nor complete in scope, but the 45 or so are far enough along that we wished to present them to the National Reading Conference members for their information and reaction.

JOSEPHINE GOLDSMITH
John Stevens High School:
Edison, New Jersey

# DIRECT EXPERIENCE READING TASKS: PSYCHOLOGICAL DEVELOPMENT[1]

The fields of psychology and reading offer complementary insights into children's functioning. While the reading teacher is most deeply and committedly a specialist, he must be aware of the goods and services offered in other shops. The aims of the psychological development section of the reading experience tasks are threefold. We want to promote easier and more productive exchange of information with school psychologists. We would like reading teachers to develop enough familiarity with theories of learning so that they might be encouraged to experiment with their application. Perhaps the most crucial objective is to spread enough knowledge about these areas to nurture the habit of straight questioning: that invaluable disinclination to be mollified or silenced by important sounding polysyllabic terms.

The tasks in this area range from a base in physiological psychology with experiments in auditory and visual acuity and perception through developmental tests such as Draw-A-Man to verbal and pictorial projectives. Some of the theoretical experiments cover areas such as behavior modification, learning modality, attention span and learning set. There are also experiments on the development of an interest inventory and the use of a mnemonic device.

---

[1] This is paper two (of five) from the symposium, Training reading teachers without lecture or textbooks.

### *Behavior Modification*

The experiment in behavior modification introduces the student to some of the basic processes of classical and operant conditioning. Among the concepts introduced are timing and schedules of reinforcement, positive, and negative reinforcement, successive approximation and secondary conditioned reinforcers. The student is then given a chance to try his hand at behavior modification through the use of a symbol reading task taught, with and without rewards, to a primary school child. The student is asked to draw conclusions about the efficacy of reinforcement. References to Deese and Hulse and to Staats and Staats are provided.

### *Projective Tests: Verbal Stimuli*

The experiment on projective testing, verbal stimuli gives a brief background on the history and uses of this type of projective instrument. Categories of responses considered popular and those indicative of strong affective reaction are deliniated. The student is cautioned that the diagnostic use of projectives is confined to the clinical psychologist and is asked to consider the use of informal versions to tap student attitudes in the classroom. The student is given one word association test and one sentence completion test to administer and is asked to construct a ten item inventory yielding information about student attitudes toward reading.

These tasks, planned as supplementary to classroom instruction, are built around a standardized format with headings covering purpose, materials, subjects, working time, background and a lab sheet to be returned to the instructor with the student's findings. They are designed to expose the student to a wide variety of psychological issues: the controversial as well as the crucial.

JOSEPH V. DePIERRO
Seton Hall University

# DIRECT EXPERIENCE READING TASKS: SYNTAX, WORD PART DELETIONS AND LINGUISTIC TERMINOLOGY[1]

Within this section, the three task areas to be discussed focus on basic concepts of language structure and decoding skills.

Upon completion of the syntactic tasks the student should understand that the English language has a basic structure, that there are fundamental relationships which exist within the structures underlying sentences and that these basic structures and relationships are vitally involved in communication and, hence, comprehension of the printed page.

### *Generative Grammar*

Three of the tasks in this category fall under the heading, "Generative Grammar" and are simply labeled Exercises I, II, and III. Exercise I is concerned with grammatical and semantic features and involves the student with nonsense sentences such as "The flumy flutzels fuppled fursily in the furish fud." Morphemes which carry grammatical relationships are retained while all others are replaced by nonsense structures. The student reads the nonsense sentence and is asked questions such as "What kind of flutzel were they?" or "What did the flutzels do?" Exercise II gives the student practice in sentence diagramming according to the "tree" patterns of generative grammarians. Exercise III deals with transforms and the process of embedding as methods of developing complex syntactic patterns.

---

[1] This is paper three (of five) from the symposium, Training reading teachers without lecture or textbooks.

### *Word Class Deletion*

Word Class Deletion, another area within syntactic tasks, tries to determine if certain classes of words are more vital to comprehension than others. In turn, specific word classes such as nouns, modifiers, verbs and others are deleted in cloze fashion from paragraphs. The student is directed to replace the missing words. After checking his answers he notes which classes of words were most difficult to replace and discusses why.

### *Word Part Deletion*

The basic purpose of the Word Part Deletions, the second task category, is to determine if certain word parts are more important than others in recognizing words. The student is presented with eight columns of sentences and words within which specific word parts have been omitted. The student is directed to recognize each semi-deleted word and to record and check his responses. For example, Column A consists of five sentences, each of which contains a word with all vowels deleted. The student must read each sentence and try to recognize the words with the deleted vowels. In contrast to this Column B contains sentences within which are words with all consonants deleted. Other "deletion pairs" such as bottom half vs. top half, first half vs. last half, and initial letter vs. final letter are also utilized. Upon completion of all exercises the student notes which deletions were most detrimental to word recognition.

### *Linguistic Terminology*

The task entitled Linguistic Terminology teaches the student to understand the terms phoneme, grapheme, allophone and morpheme in rudimentary form. The task accomplishes this through several mini-explanations and exercises such as phoneme-grapheme counts and comparisons and morpheme counts. Distinctions are made between such concepts as source and system morphemes and free and bound morphemes.

MARY M. BYRD
Westfield Public Schools
Westfield, N. J.

# DIRECT EXPERIENCE READING TASKS: READING MEASUREMENT[1]

The tasks which have been placed in this section are not limited to the measurement of reading skill, but include tasks designed for several purposes. One purpose is the demonstration of the existence and relative importance of some aspects of the reading process.

## ASSORTED ASSESSMENT

### *Eye-Voice Span*

Eye-voice span is defined as the space between the word being vocalized and the word being fixed by the eye. The child is asked to read aloud from a page; as he does so, the teacher suddenly covers the page with a sheet of paper. The number of words the child continues to say after the print has been covered constitutes his eye-voice span. In general, good readers have a larger span than do poor readers. This raises some interesting questions about very complex subjects, such as the time needed for mental processing.

### *Attention Span*

Some of the tasks are designed to demonstrate variations which can be expected within any one pupil—in performance of the same task from time to time, and in performance of different but related tasks. Testing the "Attention Span" of a single subject on a simple task which is repeated over a period of time will illustrate to the student the variance in the pattern of performance. When the scores are charted, we find a similarity to the normal curve.

---

[1] This is paper four (of five) from the symposium, Training reading teachers without lecture or textbooks.

### *Capacity Determiners*

Tasks have been included to provide practice in determining various levels of pupil abilities and skills. Pupil potential may be measured by using the listening capacity level determiner. Several formulas traditionally used to estimate the level at which the pupil has the potential to perform are included in a measurement exercise. When age and IQ are known, the student is given experience in calculating the CAGP (chronological age grade placement), MAGP (mental age grade placement), XA (expectancy age—potential reading age), and XAGP (potential reading grade).

## READING LEVEL ASSESSMENT

An "Oral Reading Word List" is used as a simple method of judging a child's reading ability. The child is asked to read aloud from a list of words that increase in difficulty. Another method of determining reading ability is demonstrated by having the student administer some "Oral Reading Paragraphs" to pupils on different grade levels.

### *Test Score Interpretation*

Standardized test scores are useless if they are not properly interpreted. For this reason, exercises in interpreting these scores have been included under "Test Score Interpretation." Students find percentiles, standard scores, grade scores, and orangutan scores. They are asked to interpret these scores for pupils in terms of grade level and chance level.

### *Listening-Reading Tasks*

In some instances, similar tasks are used for different purposes. The materials used for testing "Listening vs. Reading Comprehension" and for "Determining Capacity Level" may be the same. However, the purposes behind these tasks are different, and the methods used differ somewhat. The "Listening vs. Reading Comprehension" task is used to help the student discover that there is a difference between what the pupils is able to comprehend when he is read to and what he is able to comprehend when he has to decode and organize the material himself. In this task, the student reads one passage to the subject and then asks him questions on the content. The subject is then given a passage of equal difficulty to read silently. He is asked to answer questions orally on the passage read. The results of the two sets are then compared by the student.

JOSEPH P. SIKORYAK
Middlesex County College

# DIRECT EXPERIENCE READING TASKS: COMPREHENSION AND STUDY SKILLS

The experiments explained here deal with the issues of context clues, comprehension question variations, structural organization, and reading-study systems.

## Context Clues

In the experiment on context clues, the subjects are asked to define difficult words and to circle those which they don't know. Then, they are exposed to sentences with the words used in context and asked to go back to write in the meanings of those missed. Of the ones marked originally, the reading of the sentences confirms their choice; of the ones not marked, the reading invariably provides them with the definition. In both situations the experiment heightens their sensitivity to this comprehension aid. An extension of this activity would be to have the subjects select words at the grade levels at which they intend to be teaching and create similar exercises.

## Comprehension Question Variations

In the experiment on comprehension question variations, the subjects are asked to read a passage and answer the questions which follow. In addition, a list of question models, both low level and of more sophisticated types are provided. Subjects are asked to scrutinize each question to identify which type it is. After checking the answer key, the students might be asked to select a short passage and formu-

---

[1] This is paper five (of five) from the symposium, Training reading teachers without lecture or textbooks.

late questions of similar types. Hopefully, this will encourage teachers using materials with limited question types to supplement them with questions which tap higher levels of cognitive functioning.

### STRUCTURAL ORGANIZATION

The experiment on structural organization aims at building awareness of basic forms of discourse. The subjects are asked to read a selection of about four hundred words and record the main ideas covered. Then the subjects are exposed to the answers and the methods by which they are developed in the selection. Armed with the knowledge of the organizational structure, the subjects read a second selection of similar length and are asked to record the main ideas. After checking the second set of answers and reviewing the organizational pattern, they proceed to the third selection and perform the same tasks. Hopefully, through positive test results and introspection, the subjects find that their knowledge of the structural organization made them more aggressive readers, thus contributing to greater comprehension, improved retention and increased reading rate.

### READING-STUDY SYSTEMS

The experiment on reading-study systems aims at building familiarity with various systems (SQ3R, OK4R, PQRST, etc.), so that they may be used naturally when relevant. Robinson's SQ3R is used as a model since it was the first formal system presented. After each of the elements is explained, the subjects are given five other mnemonics which are used to symbolize other published reading-study systems and asked to determine what the other letters stand for in OK4R, RSVP, PQRST, OSCAR, and OARWET. If the subjects understand what should be done first, last, and in-between, then they should be able to get a majority of the answers. The answers key would then provide them with the reinforcement needed.

These are the kinds of activities which are used to expose prospective teachers to comprehension and study skills tasks. Their purposes, as well as those mentioned in other sections, are the same: to provide the student teachers with a bit of reality in doing reading tasks which may be more effective than traditional course work.

HARRY SINGER
University of California, Riverside

# READING CONTENT SPECIALISTS FOR THE JUNIOR HIGH SCHOOL LEVEL[1,2]

Junior high school teachers have long recognized that reading disabilities can—and usually do—impair a student's learning in the content area. Unfortunately, recognition of the problem implies something considerably short of a solution. The junior high school in California is an "in-between" administrative unit and enrolls students who are at an "in-between" stage of development. Teachers are not specifically prepared to instruct this age level, which is highly varied in ability and achievement, undergoing rapid physical and physiological changes associated with puberty, and beginning to search for and establish identities, vocational goals, heterosexual relationships, and greater independence from family, yet needing and wanting emotional support and well-defined boundaries of behavior.

Despite the complexities of teaching and the educational needs of students at the junior high school level, there is no teaching credential specifically required for the junior high school level. Under the present Education Code in California, the elementary and secondary credentials completely overlap at the junior high school and in this overlap teachers are authorized to teach in the junior high schools.

---

[1] A joint project of the University of California, Riverside and the Riverside and San Bernardino Unified School Districts. This project was supported by the U.S. Office of Education's EPDA Program. Project Officer: Miss Helen O'Leary.

[2] Appreciation is expressed to my statistical research assistant, Sherrel Beasley Haight, for processing these data and to my research assistant, Grace Van Doren, for coordinating the collection of the data, and to Marka Adams for typing the final report.

The elementary credential's authorization is from grades K-9 and the secondary credential from 7-12. Actually, most junior high school teachers in the 20,000 square mile area of Riverside and San Bernardino Counties served by the University of California, Riverside hold secondary credentials. This means that very few, if any, of these teachers have had any preparation in teaching reading, either developmental, remedial, or reading in the content areas. Yet it is well known that the achievement range of junior high pupils is quite wide. Indeed, a typical class of seventh graders has a range of achievement in reading extending from the third to the eleventh grade, with approximatley half of the students achieving above and the other half below median grade level (Bond and Tinker, 1967).

A large percentage of students who are achieving below grade level are minority students. As shown by the Riverside School Study (Singer, 1970), there is an increasing grade-equivalent disparity in the reading achievement of Negro and Mexican-American ethnic groups compared with the Anglo majority at successive grade levels. By the sixth grade, the discrepancy is about two grade level equivalents.

While the causes for these discrepancies are multiple, including bilingual and dialectical handicaps (Bailey, 1966), preschool preparation (Bereiter and Engelmann, 1966), normalization of instructional expectations in *de-facto* segregated schools (Wilson, 1963), among other factors, there is some probability that at the junior high school level the initial causes for some of these deficiencies are no longer present (Singer, 1956).

However, the discrepancies are likely to increase during the junior high school years rather than be ameliorated for several reasons, including tracking practices at the junior high school level (Balow, 1964) and the inadequate preparation of teachers at this level for overcoming the deficiencies.

Instead of *de-facto* integration alleviating the problem, there are some possibilities of its exacerbating the discrepancy. In their report, Singer and Hendrick (1967) state the problem thusly:

> The integration plan in Riverside can be conceptualized as an experimental alternative to the deleterious effects of segregated education. For example, a particularly pertinent study reported by Alan Wilson in 1963 revealed that because teachers tend to normalize or base grading standards upon the average of one particular group taught, minority children in de-facto segregated elementary schools receive grades higher than their standardized test scores would suggest they should, while the converse is characteristic of educationally advantaged children. When both groups are brought together in integrated junior and senior high schools, grading standards may still be normalized, but, owing to the more representative range of achievement scores in the combined groups, assigned grades would

> then more closely approximate test scores. Consequently, on the average, relatively lower grades would probably be assigned to the minority children and higher grades to majority children. This change in relative rank in achievement and the apparent discontinuity in assigned grades in consonance with other factors could result in the minority children from de-facto segregated schools experiencing a negative cycle of lowered grades, a more competitive and unhappy environment, lowered morale and motivation to learn, and subsequent reduction in achievement. It would, therefore, not be surprising to find that these children and their parents would conclude not only that their elementary schools were inferior, but that education in junior and senior high schools is both inappropriate and not responsive to their needs. Tracking systems at junior and senior high schools which use differentiated curricula can only compound the problem because they tend to increase the disparity in achievement among children in the various tracks (p. 146).

While the causes for students dropping out of school are also complex and multiple, one of them is frustration experienced by students who are intelligent but whose achievements are far below their capability level. At the present time in Riverside some students tend to drop out in the eleventh grade when school attendance is no longer compulsory because their achievement deficiencies, particularly in reading, are embarrassing to them. Therefore, one instructional need is preparation of teachers for the junior high school and specifically for teaching minority children who have recently been *de-facto* integrated.

A second instructional need at the junior high level is teaching students to read in such content areas as science, history, and social studies. At the present time, students are expected to transfer skills acquired in elementary school in developmental reading, based on literary content, to all the other content areas. While there are some communalities, there are specific differences in the content areas that can and should be taught. For example, students need to learn how to read and interpret charts, maps, graphs, and tables. They need to learn various types of reading, such as reference and critical reading (Gans, 1940). They need to be aware of various patterns of writing such as problem-solution, cause and effect, question and answer, inductive and deductive, and a specific technical vocabulary, know available reference works, and learn different reading techniques, such as analytical techniques for critical reading and skimming techniques for overviews. They also need to develop a variety of rates and learn to use them appropriately. Thus, reading in the content areas includes knowledge of technical vocabulary, interpretation of data, unique patterns of mental organization required to read the printed materials in each content field, devices, symbols, illustrations, and specialized purposes, motives, and attitudes, which do not usually occur in tests of general reading ability (Russell and Fea, 1963; Smith, 1965; Moore, 1969; Singer, 1969). Since secondary credential requirements do not include any specific instruction in teaching any type fo reading in any content area, most junior high school teachers need such instruction and it is likely that such preparation will probably be reflected in pupil achievement in the content areas.

To teach reading in the content areas to students whose achievement covers an eight to ten year range requires not only knowledge of the psychology and pedagogy of reading, but also more advanced knowledge than the typical teacher possesses of the structure of subject matter in the various content areas (Bruner, 1963). Consequently, half of the graduate work in the program consisted of work taken in an ancillary field, e.g., history, English, psychology. Such knowledge should enable the teacher to restructure the content of the curriculum and identify appropriate instructional objectives, diagnose content deficiencies, and select specific material for remedying gaps in students' knowledge (Gagné, 1966) and for promoting greater teaching efficiency and effectiveness. Graduates of the program should therefore be able to provide leadership in the public schools in the evaluation and reorganization of curriculum and improvement of teaching and learning in the subject fields. (Decentralization of the administrative structure, occurring in Riverside as well as elsewhere, requires more qualified teachers in the schools to fulfill functions formerly conducted by central administration consultants and supervisors.) They should also be able to bring together and relate the content and methods of ancillary fields to the problems of human learning in the context of the school.

## PROJECT PROGRAM

The project consisted of a two year program. The first year involved concentrated study at the University with field work experience in nearby junior high schools. The second year of the program was an intern type of experience in junior high schools, two in Riverside and two in San Bernardino Unified School Districts.

During the first year, students took approximately five courses in Education and five courses in an ancillary department. The following were the required courses in Education: Literature and Reading in the Secondary Schools; Theory, Diagnosis and Treatment of Reading Difficulties; Curriculum and Instruction; Review of Research in Literature in Education; and Seminar in Sociology of Education. In the ancillary departments, candidates had choices to fit their interests and academic backgrounds. Because of self-selection, screening, and final selection criteria, eight candidates were in English, two in history, one in psychology, and one in Spanish. The candidates ranged in age from 23 to 50, nine were female and three male, ten were from California and one from Minnesota and one from Pennsylvania. The range in teaching experience was from two to nine years. Two candidates were teaching at the elementary level and the other ten were at the junior or senior high school level when they entered the program.

*Specific Objectives of the Program*

1. To improve the reading achievement of junior high students, particularly those who are educationally disadvantaged. The intent was to attain statistical convergence in the achievement of Anglo and ethnic minority groups and, in general, to improve the average level of reading achievement in the junior high school.

2. To prepare reading specialists for the junior high school. These specialists were to learn to diagnose and remedy reading difficulties through course work and a year-long internship under supervision.

3. To spread the effects of the reading specialist to all teachers in the junior high school through consultative and in-service training in teaching reading in the content areas. The purpose of the consultation with individual teachers was to ameliorate the reading abilities of students, and to have the classroom teacher become aware of reading difficulties in the content areas and ways of overcoming such difficulties. The teacher could then incorporate such knowledge into classroom practices that would be beneficial to all students in the class.

4. To improve the junior high school curriculum. The program presumed that improved competence in the subject field would increase the teacher's awareness of instructional options while he was engaged in the process of curriculum decision-making. Thus the course work in a specific content area was intended to work as a vehicle for helping the teacher improve the curriculum in his field.

5. To develop model classrooms in the junior high school. It was expected that each of the content reading specialists would develop a model classroom for the improvement of reading in his respective content areas, e.g., science, social studies, English, etc. Development of these classrooms was viewed as a central objective of the entire project.

*Junior High Schools*

The four junior high schools were in Riverside (two) and San Bernardino (two). Principals had volunteered their schools. In San Bernardino the program was installed in two brand new schools.

A one day conference of department chairman in science, social studies, mathematics, and English, and principals of the four schools was called in May of the first year. Dr. Harold Herber, Syracuse University, taught the participants the use of reading and reasoning guides and showed films of their use in the classroom. In the afternoon ,the Content Reading Specialists served as consultants to each content group in the art of constructing reading and reasoning guides to accompany textbook assignments and teach students processes and content in the assigned material. This conference set the stage for the second year when the Content Reading Specialists would serve as teachers and teacher-consultants in schools represented at the Conference. It also initiated and gave confidence to the Content Reading Specialists in their roles as consultants.

## GENERAL PROCEDURES

Four major approaches were used in teaching reading in the content areas to heterogeneous classes of seventh graders and a few classes of eighth graders. These approaches were: the project or theme method of teaching a unit, e.g., the Scholastic Magazines type of units of "Courage" is an example of the project method; Herber's (1968) reading and reasoning guides were emphasized in each content area with an extension made by our program to include word recognition and word meaning instruction; Moffett's (1969) techniques were stressed in English classes: group discussion, reading taught through writing activities, journals were kept, paperbacks were used in self-selection and free reading time, role playing and class plays were also utilized; and cross ability teaching was employed, sometimes in group settings and other times in one-to-one relationships.

The Content Reading Specialists also spent one-third of their day consulting with other teachers. This consultation varied from teaching the use of study guides, to working with small groups, to working with and demonstrating reading instruction in the content areas to an entire class and its teacher. To encourage teachers to consult and to provide the means for curriculum modification, each junior high school was given $2500 for purchase of materials and supplies. Purchase decisions were to grow out of consultation between teachers and the Reading Content Specialists. These materials are now housed in the University's library where they are availale for loan to any teacher in the University's service area. The success of this feature of the project was attested to by the incoming president of the San Bernardino teaching association who had consulted extensively and was committed to spreading the effects of the program throughout San Bernardino.

The leadership objective of the program was also realized for upon completion of the project, half the graduates immediately gained positions related to their training and experience. Anthony Bechtold became a reading consultant in Poway School District, Poway, California; Richard Zimmerman is now a reading specialist at Moreno Valley High School, Sunnymead, California; Christina Guitierrez and Jean Fruehan became chairman of their respective social studies and English departments in Riverside and San Bernardino junior high schools; Margaret Minor was employed in a newly created position as the reading consultant in the Riverside junior high school where she had previously taught. Bonnie Bauman returned to her high school and Vaughan Hudson costinued on in her junior high as an English teacher. Dale Johnson and David Kahl, on sabbatical leave from their high schools in Oceanside, California and Erie, Pennsylvania, respectively, had to return to them. Patsy Miller took the year off from teaching. Sylvia Cherry accepted a position as part-time member of the University of California, Riverside's supervisory staff in teacher education and part-time teacher in the junior high school in Riverside where she had taught before entering the program.

## DATA

Data were gathered in the Content Specialists' classes and in control group classes. The control group classes consisted of teachers who volunteered to have their classes tested in the fall and retested eight months later in May. It was not feasible to use the most defensible design of random assignment of students to experimental and control group treatment.

Three test batteries were used: California Reading Achievement Test, Junior High School Level, Forms W and X, (19); Carter's California Study Methods Test (19); and Athey-Holmes Reading Personality Scale (1969).

## RESULTS

Analysis of covariance comparisons were made of Experimental vs. Control Groups by total, by school, by teacher, and by ethnic group on reading ,study methods, and personality. The results were not statistically significant.

Next a design was used based upon the attitudes of the Content Reading Specialists towards teaching reading in the content areas (Otto, 1969). Although the group had a narrow range of attitudes, the range nevertheless varied from 51 to 65. (According to the norms for a group of 38 junior high teachers, the mean was 41.9 and standard deviation 9.39, Otto, 1969.) Dividing this range into half, the top and bottom halves were compared against the upper and lower 27 percent of each of the ethnic groups. The result of this analysis of covariance test was also not statistically significant.

## DISCUSSION

While the results were statistically insignificant, the project did succeed in demonstrating that it was possible to teach a heterogeneous class of junior high school pupils where teachers had been used to tracking students. The consequence of this heterogeneous grouping in a school district consisting of about six percent Black, 12 percent Chicano, and 82 percent Anglo students showed no loss in reading achievement. Of course, it was disappointing that we had not at least bent the Anglo majority vs. Black and Chicano minority curves of achievement. We can only speculate that perhaps a longer period of teaching than one year is necessary to make a significant impact on group means and to overcome the deleterious effects of previous years of growing discrepancies in majority and minority group achievement.

Further and more massive attempts, reaching down into the elementary school and extending up into the high school, may be necessary to effect a change in group differences in achievement. If schools are to carry out the 1954 Supreme Court definition of equality of educational opportunity, then more differentiated input is required. As interpreted by Coleman (1968), the criterion for equality of educational opportunity was shifted by the Supreme Court's decision from equal input to equal output. In effect, the concept under which the education of exceptional children has long operated has been generalized by the Supreme Court to include minority groups (Singer, in preparation). For moral, legal, political, and educational reasons, school districts such as Riverside terminated *de-facto* segregation some six years ago and active attempts have been made and are still being made by various agencies, including federally funded programs such as the present one, for the investigation and amelioration of the achievement and adjustment of minority children, yet the evidence indicates the inequalities in achievement between majority and minority groups still remain (Singer and Hendrick, 1967; Singer, 1970). Thus, some 17 years since the concept of equality of educational opportunity has undergone its revolutionary change, its attainment continues to challenge our imagination and ingenuity to transform into reality a dream of a promised land.

## REFERENCES

Athey, I. J., & Holmes, J. S. Reading success and personality characteristics in junior high school students. Publications in Education No. 18. Berkeley: University of California, 1969.

Bailey, B. Report to the meeting of the National Council of Teachers of English, 1965. Summarized in C. Bereiter & S. Englemann, *Teaching disadvantaged children in the preschool*. Englewood Cliffs: Prentice-Hall, 1966. P. 45.

Balow, I. H. The effects of homogeneous grouping in seventh grade arithmatic. *The Arithmetic Teacher*. 1964, *2*, 186-191.

Bereiter, C., & Englemann, S. *Teaching disadvantaged children in the preschool*. Englewood Cliffs: Prentice-Hall, 1966.

Bond, G., & Tinker, M. *Reading difficulties*. New York: Appleton-Century-Crofts, 1967.

Bruner, J. *The process of education*. New York: Vintage Book, 1963.

Gagne, R. Contributions of learning to human development. Unpublished address of the Vice-president, Section I—Psychology—American Association for the Advancement of Science, Washington, D. C., 1966.

Herber, H. L. *Reading in the content areas*. Englewood Cliffs: Prentice-Hall, 1968.

Moore, W. J. About reading in the content fields?*English Journal*, 1969, 707-718.

Moffett, J. A student-centered language arts curriculum, grades K-13: a handbook for teachers. New York: Houghton Mifflin, 1968.

Otto, W. Junior and senior high school teachers' attitudes toward teaching reading in the content areas. In G. B. Schick & M. M. May (Eds.), *The psychology of reading behavior*. National Reading Conference 18th Yearbook. Milwaukee: the Conference, 1969, Pp. 49-54.

Russell, D., & Fea, H. Research on teaching reading. In N. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNalley, 1963. Pp. 865-928.

Singer, H. Bilingualism in elementary education. *Modern Language Journal*, 1956, *40*, 444-458.

Singer, H. Factors involved in general reading ability and reading in the context areas. In G. B. Schick & M. M. May (Eds.), *Reading process and pedagogy. National Reading Conference 19th Yearbook*. (Vol. 1) Milwaukee: the Conference, 1970. Pp. 295-305.

Singer, H. Effect of achievement in anglos, blacks, and mexican-americans. *American Educational Research Association Abstracts*, 1970.

Singer, H. Effort of integration on achievement of anglos, blacks, and chicanos—a five year longitudinal study. In H. Gerard & R. Jones (Eds.), *Journal of Social Issues*, anticipated date: Jan. 1973. (In preparation.)

Singer, H., & Hendrick, I. G. Total school integration: an experiment in social reconstruction. Phi Delta Kappan, 1967, *49*, 143-147.

Smith, N. B. Reading in subject matter fields. *Educational Leadership*, March 1965, 382-385.

Wilson, A. B. Educational consequences of segregation in a California community. In Report of the U. S. Commission on Civil Rights, *Racial isolation in the public schools*. (Vol. 2) Washington, D. C.: Superintendent of Documents, CRI, 2: Sch 6/12/V.2., 1967. See also, A. B. Wilson, Social stratification and academic achievement. In H. Passow (Ed.), *Education in depressed areas*. New York: Teachers College Press, 1963. Pp. 217-235.

EDMUND ZAZZERA
Educational Developmental Laboratories

# MODEL OF THE DYNAMIC ACTIVITY OF READING AND THE READING TEACHER

The focus of my talk is a model of the reading process developed at Educational Developmental Laboratories by Stanford E. Taylor (1971). The model is fully described in its present form in an EDL Research and Information Bulletin. Within the scope of this paper, I cannot hope to describe the entire model, nor develop fully the articulation of its parts, but I can tell you why it was developed, what purpose it serves at present, and what we see for it in the future.

## THE DEVELOPMENT OF THE READING PROCESS MODEL

The model was originally developed for use in teacher education workshops conducted at and by EDL. Its purpose was to help reading teachers to appreciate the intricacies of the reading process. This purpose sets it apart from most, if not all, other models in that it was developed to serve on a utilitarian and practical level as opposed to a pure research function.

The model traces the reading process from the stimulus through to meaning and beyond. Here we see the major processes and all the subfunctions or factors that affect them.

### *Four Main Divisions of the Model*

*Seeing* refers to the purely physiological stage during which the stimulus, or the light reflected from around the print, is received by the eye and transmitted to the brain.

*Perceiving* involves psychological and physiological activity. Here the retinal pattern is transformed into an identified or recognized

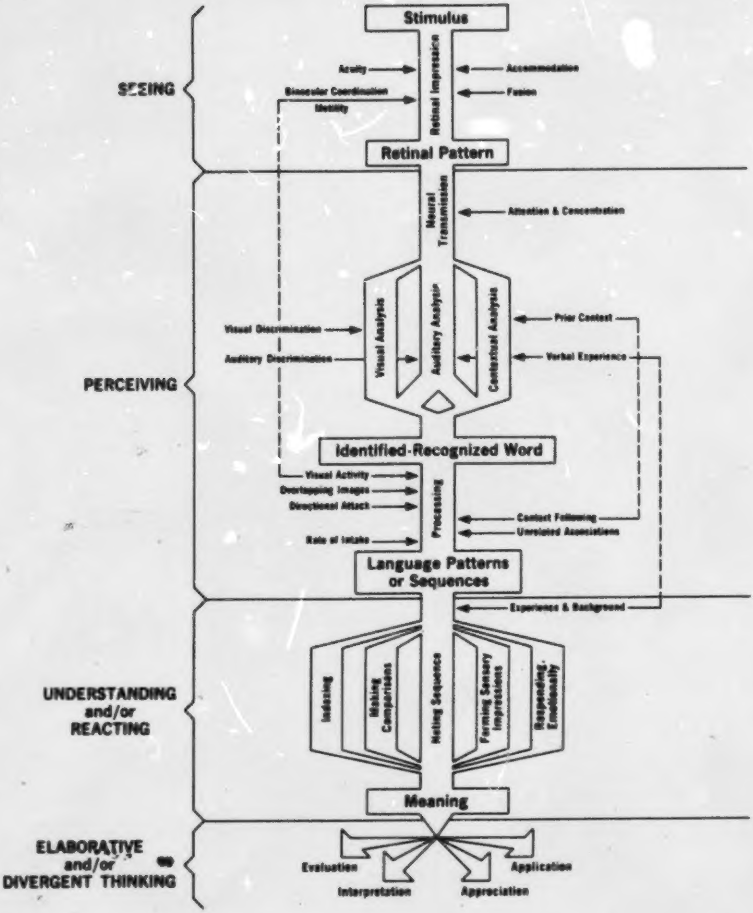

SEEING
PERCEIVING
UNDERSTANDING and/or REACTING
ELABORATIVE and/or DIVERGENT THINKING
Stimulus
Retinal Impression
Acuity
Accommodation
Binocular Coordination
Motility
Fusion
Retinal Pattern
Neural Transmission
Attention & Concentration
Visual Analysis
Auditory Analysis
Contextual Analysis
Visual Discrimination
Auditory Discrimination
Prior Context
Verbal Experience
Identified-Recognized Word
Visual Activity
Overlapping Images
Directional Attack
Rate of Intake
Processing
Context Following
Unrelated Associations
Language Patterns or Sequences
Experience & Background
Indexing
Making Comparisons
Noting Sequence
Forming Sensory Impressions
Responding Emotionally
Meaning
Evaluation
Interpretation
Appreciation
Application

Fig. 1. Reading Process Model

word and then words are processed into language patterns or sequences.

*Understanding and/or Reacting* refers to that stage where word sequences are accepted in light of the reader's experience and background and are translated into meaning.

Once meaning is established, the reader engages in *Elaborative and/or Divergent Thinking* as he evaluates, interprets, appreciates, or applies what he has read.

Obviously reading does not take place in such a simplistic or linear fashion. The process is more nearly cyclical, with fresh retinal patterns being received while "old" patterns are being processed into thought units. Why are we using two dimensions to represent what is really a four-dimensional process? The answer goes back to the utilitarian nature of the model.

In order to reach many people, it had to meet two criteria. First, it had to be in printable form, which necessitates two dimensions. A true representation would have to include the time function and, therefore, require a more sophisticated media for presentation. Second, the model had to be easy to understand, or stated in a form and in terms that could be quickly communicated to people who do not have an extensive background in learning theory.

## Overview of the Model

Bearing in mind the limitations of this two-dimensional representation, let us take a look at the process.

### *Seeing*

Seeing refers to the stage in which the stimulus is transformed into a retinal pattern which is dispatched to the brain. More simply put, this is the reception stage. The reader's success or failure at this stage depends on the legibility of the stimulus that is presented to him and how well his visual acuity, accommodation, fusion, binocular coordination, and motility allow him to function in receiving the stmulus.

### *Perceiving*

Now we enter the second major stage, perceiving. Here the retinal pattern is transmitted to the brain. How well this is accomplished depends on the attention that the reader can focus on the task and he concenration he can sustain.

Three kinds of analysis allow the neural transmission to be received as a word. There is varying stress placed on each of these functions at different stages of the reader's development. With mature readers, visual analysis is carried on subliminally, and reinforced by contextual analysis. Auditory analysis, or souding out words, is used mainly by less skilled readers.

The reader's skill in using these forms of analysis depends on his visual and auditory discrimination, his ability to use context clues (both prior context and context following), and his verbal experience.

To trace the events of this stage, it seems probable that a reader uses prior context in order to anticipate both words and ideas, employs visual analysis upon the instant of contact with a word in order to effect identification and recognition or storage, sometimes deviates from reading to employ auditory analysis in order to unlock and recognize less familiar words, and then uses the "context following" in orde rto confirm the accuracy of specific words selected. Recognition then must be thought of, for the most part, as a multi-faceted process that typically lags behind visual intake.

The next major stage is processing. Processing can be defined as a stage during which words are stored and sufficiently stabilized in a manner that will allow the mto be retrieved along with other words and realized as language patterns or sequences. The stage of processing is maintained as an ongoing activity concurrent with the process of recognition, but lagging slightly behind.

The effectiveness of this stage depends on the reader's visual activity, his ability to cope with overlapping images, the orderliness of his directional attack, the effect of unrelated associations, and his ability to control his rate of intake. Thus, as each word is processed, it is added to and becomes part of a larger message which is recognized in the form of language patterns.

### *Understanding and/or Reacting*

Now the reader will understand or react to an evergrowing and accumulating message. This message passes through one or more cognitive or affective processes and is realized as meaning. These processes are more often used in combination than singly. The success of the reader's efforts at this stage depends largely on the quality and quantity of his experience and background, or what has been called his mental content.

In indexing, the reader assigns relative values to bits of information, creating a mental outline. The reader makes comparisons, noting similarities and differences, sorting the known from the unknown. In notig sequece, the reader arrages material according to time, space, degree, etc. The reader can also react with his senses and/or respond emotionally.

### *Elaborative and/or Divergent Thinking*

More mature readers go beyond the realization of meaning to engage in various forms of elaborative and divergent thinking. During this stage, he evaluates what he has read, recognizing whether the events are natural developments or implausible, whether the characters are real or stereotyped, whether the tone is sincere or artificial. He proceeds to make additional associations and interpretations, relating what he has read to his own experiences, convictions, and interests. He might appreciate the esthetic content of a message and its form, as he pauses to reflect on the quality and nature of the reading material he has experienced. Or he may apply what he has read: creating something new, taking direct action, or altering his manner of behavior.

### SIGNIFICANCE OF THE MODEL TO THE READING TEACHER

What does this model mean to the reading teacher? More important, how can he or she use it?

First, it is important for teachers to realize that the process of reading consists of all these factors and subfunctions. A program of instruction should not be considered complete unless it takes into account all of the factors described. Thus the first purpose of the model is to serve as a checklist for evaluating the comprehensiveness of a reading program. Does it include techniques for dealing with all vital functions and subfunctions? A program may be underweight in a number of areas, but it can also be overweight. For example, some secondary programs over-stress speed reading, while others devote all of the students' time to skills of understanding content.

More specifically, in the seeing section, most teachers are cognizant of the importance of visual acuity in the reading process. However, teachers should also be aware of the effect of accommodation, fusion, binocular coordination, and motility on the retinal impression. This awareness will enable them to evaluate the visual screening program in use in their schools and to insist on a more comprehensive program ,if necessary. Teachers should also know that training techniques exist today which make possible improvements in ocular motility and accommodation. Similarly, there are specific instructional techniques for training attention and concentration, and for enabling students to cope with overlapping images, for example.

In the past, reading has been treated classically as a thinking process, so most reading programs are well equipped to provide comprehensive instruction in the areas of understanding and/or reacting, and to guide students in making meaningful use of reading experiences through evaluation, interpretation, appreciation, and application. However, it is important to place these last two areas in proper prospective.

### FUTURE OF THE READING PROCESS MODEL

A final question: What is the future of this model of the reading process? It is being published now with the understanding that it has not been thoroughly researched and documented, but with the intention that this publication can serve as a base line for future development. Indeed, with its publication, EDL/McGraw-Hill is initiating a research program to refine the model and further explore its intricacies. It is our hope that people such as yourselves will be interested in working on various types of projects that will study all the functions of the process. We would like to see this model compared with Goodman's reading process model and the Guilford's model of the intellect, and checked out against Bloom's taxonomy, just to mentio[illegible] a few areas of potential exploration. Various facets of the model nee[illegible] to be explored in depth. Further, the interface between the various stages of the model and instructional technology should be explored. The end product of this research program should be a model that will be useful not only for researchers in coming to a deeper understanding of the fascinating process of reading, but also for teachers who are faced daily with the responsibility for helping students to become more efficient and more effective readers.

## REFERENCES

Taylor, S. E., *The Dynamic Activity of Reading: A Model of the Process*, Research and Infirmation Bulletin No. 9, Huntington, N. Y.: Educational Developmental Laboritories, 1971.

EMERY P. BLIESMER
Pennsylvania State University

# 1971 REVIEW OF RESEARCH ON COLLEGE-ADULT READING

The reports cited in this present review are, in general, ones which have appeared (and were available to this reviewer) during the period of September, 1970, through September, 1971. The number of reports included as relevant for this year's review, 192, is slightly more than the number included in last year's review (Bliesmer, 1971). The Yearbook for the 1969 National Reading Conference (Schick & May, 1970a&b) was not available in time for those of its contents which were pertinent to be included last year; so it has been included in the present review. The Yearbook for the 1970 Conference (Greene, 1971) became available just as final preparations for this review were being made. It was possible to include pertinent references from the last Yearbook only because the number of reports found in the volume which were relevant for the areas of college-adult reading research was considerably less than the number found pertinent in the Yearbook for each of a number of years previously. The *Proceedings* of the Western College Reading Association's fourth annual conference (Christ, 1971) was received too late for inclusion of specific articles in this review. The contents appeared to be mainly descriptions of specific programs in the WCRA area; but several included evaluation reports or other research aspects (e.g., presentations by Adams, Biggs and others, Hagstrom, and Reid). The North Central Reading Association's *Sixth Yearbook* (Wark, 1971a) also contained several relevant reports.

Consistent with past reviews, the references included in this one have been restricted in general to those which deal with, or have fairly direct relevance to reading at college and other adult levels. The review has again been loosely organized or grouped under the several major divisions indicated.

## PROGRAMS

### *Surveys*

Several surveys dealing with college and other adult reading programs have again been reported. Among conclusions reached by Dare (1971), based on results obtained with questionnaires sent to 559 two-year colleges, were that nearly all public two-year colleges provided some type of remedial instruction (English and mathematics as well as reading); most colleges refrained from using the term remedial and used such terms as developmental or guided study instead; while preliminary studies preceded the establishment of most programs, few colleges made studies of the programs once they were started (but attitude of faculty and students was reported as favorable in most colleges); and purposes for remedial programs were considered by most colleges to be "salvage, second-chance and redirection". Lowe and Stefurak (1970) presented results from another state survey (Georgia) of college reading improvement programs. Of the 30 out of 58 schools replying, 9 reported that they were not offering reading improvement programs. Franklin (1971) reported results from a questionnaire survey of diagnostic procedures in college and university reading clinics in the United States. (Questionnaires were mailed to 741 clinics; but the percent of response was not indicated.)

Among findings obtained by Berger (1970) through the mailing of questionnaires to corporations, commercial reading firms, and college and university reading centers in the U.S. and Canada (with 27% and 33% response from the U.S. and Canada, respectively) were that corporations in Canada tend to have a greater percentage of high school graduates than U. S. corporations, with a greater percentage of college graduates taking reading improvement courses; specific rate goals of programs ranged from 25 to 18,000 words per minute; respondents indicated gains were maintained; and relatively few colleges and universities provided a reading service for adults.

### *Evaluations of Programs*

Purposes for presenting a number of programs at this conference last year were mainly ones of description (Cartwright, 1970; Christ, 1970; Ewing, 1970); but some also included reports of evaluations to some extent. Kerstiens reported that in the El Cabino College program "point-percentile gains on standardized tests are substantial" and that "we can report measurable progress" on low percentage of attrition, a diminishing flunking rate, and higher passing grades (1970, p. 110). Stewart reported that the mean pre-post reading test score changes have "ranged from 2.0–3.5 grade points per semester from 1967 to date [1970, p. 206]." She also reported that, in a comparison of freshmen groups using cloze and program procedures, both groups made modest gain scores, but not significantly different from one another. Garofalo (1970) did not present any evaluation data in his report of a college program at the University of Maryland, Eastern Shore; but he commented that evaluation was difficult and that the worth of a college program ". . . should be measured in the same way that a library's worth is determined: by the number of students using the program and by the breadth of resources made available to the college community. This method of measurement is far more useful than standardized test scores, or the feelings of the instructor or the students [p. 110]."

Brown, Wehe, Zeinker, and Haslam (1970) found a group of 111 Southwest Texas State University potential dropout students who received counseling from selected and trained upper classmen were significantly superior to a matched control group (which had received no counseling) on prepost counseling SSHA and *Effective Study Test* Scores, on post-counseling study skill survey scores, and on first-semester grade-point averages. Caple (1969) found members of a study group of Oklahoma college students who were required to attend residence hall two-hour study sessions four nights a week for the last half of the semester (after mid-term grade reports) did not differ significantly in achievement at the semester's end from a matched control group who had not participated in organized group study sessions. When Keenen (1971) compared achievement. for three quarters, of groups of University of Florida freshmen who had been admitted even though they fell below minimum basic admission requirements, he found grade-point averages of groups which had received special guidance (including reading improvement programs) and sectioning and reduced course loads did not differ significantly in grade-point averages from groups which had received no special treatment.

Retesting results obtained by Payne (1971) from 126 students, out of an original 160, who had participated in the Northwest State University of Louisiana reading program showed that participants had retined and improved reading rate after 5 semesters, that a control group lost in rate the first semester but showed improvement after 5 semesters, that both experimental control groups retained and improved in vocabulary and total reading score gains, that participants successfully pursued more credit hours than did non-participating students, but that achievement in English, mathematics, social science, and natural science was little affected by participation. Pearson (1970) found no significant differences on various measured variables (including reading skills, study habits and attitudes, college achievement, and self-concept) betwen control groups and groups who had group counseling included in a reading improvement program.

Results obtained by Schneider (1971) in experimental remedial English classes in which developmental reading as well as conventional content was taught indicated that emphasis on reading skills could lead to improved writing; but she did not find significant differences over control groups in either writing or reading test score gains. Questionnaire and interview responses obtained by J. L. Smith (1970) from students who had participated in reading and study skills programs at Indiana University reportedly indicated that the program was meeting the needs of the students and presented no major operating problems.

City University of New York freshmen who received 3 hours a week of tutorial instruction, from faculty members who were participating in a federally funded four-week summer reading and writing skills institute, practically all indicated (through written evaluation) that they believed they had made improvement in vocabulary development, reading for main ideas, and study-type reading and that they had been motivated by instructor personnel concerned to assume initiative for further upgrading their reading and study skills, according to Newman (1971).

### *Pre-College Programs*

Several reports were concerned mainly with pre-college programs. Holmes (1971) found a significant relationship between participation in a special summer college preparatory and enrichment program at Jarvis Christian College and academic achievement during the freshmen year. Johnson, Jacobson and Beaber (1970) reported that their analysis of test score gains of a group of college-bound rural high school students in a reading improvment program indicated that vocabulary and comprehension gains were related to "genotypical" variables (such as CA, I.Q., socioeconomic status, sex, race) while reading rate gains were influenced by "phenotypical" variables (such as school attended and college selected).

A group of 184 high-risk high school graduates who participated in Project Speed (Summer Program for Educational Enrichment and Development) in three Georgia junior colleges, which program included instruction in reading, writing, study, and other basic academic skills, and group and individual counseling, were found by Miller, Artenen and Duncan (1971) to score significantly higher than a control group on 11 of 14 variables on which groups had been simular at the beginning (variables including *Nelson-Denny Reading Test* Scores). Pepper (1970) reported that 155 of 182 marginal students enrolled in a Wayne State University summer reading and study skills course later completed four quarters of academic work. Among activities Opison (1970) found reported least often useful by 450 students who participated in a Pennsylvania State University preregistration counseling program during a 9-day period during the summer of 1967 was study skills information.

*Comparisons of Methods*

Blevins (1971) found a Contrived Contextual method of vocabulary instruction (isolated word study) to be inferior to Wide Reading Natural Contextual and to Practical Senior High School English programs (in which broader natural contexts of messages were conveyed by connected words) for 36 adult institute subjects participating in a thirteen-weeks program. Curry and Roberts (1970) gave a more detailed report of Roberts' dissertation study, in which it was found that none of different schedule practices of reading program groups produced reading and spelling achievement significantly superior to that of any other practice schedule group; but practice was found to be superior to no practice at all. DiSalvi (1970, 1971) found that experimental groups of Newark State College freshmen who had participated in an evening reading and study techniques program obtained higher means than did a control group on study orientation, level of comprehension, and speed of comprehension scores, but not on grade-point averages. Differences among all groups (including control) were not significant, except for speed of comprehension. Experimental group students in a Pennsylvania junior college reading program who participated in a group counseling sessions during one of the three regular weekly reading periods each week were found by Florak (1971) to have increased in ego-strength significantly but not significantly (from regular program groups) in study attitudes or in reading performance (although removing students from one regular class meeting a week from counseling did not appear to "impede" their progress). Guice (1969) found no significant differences in reading test scores between students in reading improvement programs who had regularly practiced with cloze exercises and those who had not.

Phillips (1970, 1970) found that neither the control nor any of the three instructional approaches (teacher guided, individualized, and audio-visual) used in a reading and study skills program with groups of disadvantaged black college freshmen appeared to be significantly more effective for improving level and speed of comprehension (*Davis Reading Test*), with the individualized approach being found most effective for positive modification of study habits and attitudes (SSHA). Whitehill and Rubin (1971) found that experimental groups of graduate and undergraduate University of Wisconsin students in free, voluntary, developmental reading courses who used instrument instruction (with paperback books) had significantly higher words-per-minute rate gain scores than did a control group which had received "traditional" instruction. No significant main effects were found on the *Eyzenck Personality Inventory* for the intraversion-extraversion dimension.

### *Adult Programs*

*Programs with non-college adults were involved in several other* reports also. No real evaluation of the University of British Columbia continuing education center program for community adults was reported by Ahrendt (1971); but he did state that there had been only a five percent dropout rate over a three-year period (no refunds of fee after the second meeting); and responses to the program had been interpreted as indicating "that adults want to become efficient readers and will work to achieve this goal [p. 10]." Brooks (1970) described the origin, in 1968, and growth of classes for adults in remedial writing and reading in England and methods used and some problems encountered. He stated that the effects of the program were "impossible to measure" but that tests which were given at six-months intervals indicated "a shift of a few points in an upward direction; and that the classes seemed to have had the effect of helping many to advance socially, and also at their work [p. 155]." Inmates in two Ohio penitentiaries who used programmed textbooks in adult basic education courses were found by Carson (1971) to be able to show definite gains in reading vocabulary, comprehension, and total scores (ABE), at both easy and difficult levels. Fant's (1970) comparison group of Atlanta adult program participants who had received twelve weeks of training to improve reading skills with a Communications Skills Laboratory made significant gains in word meaning and paragraph meaning (Stanford) scores, while an experimental group who had received training with O. K. Moore's Talking Typewriter made significant gains only in paragraph meaning scores, with a comparison group having made significantly greater gains. McDaniels report (1971) of the 1970 summer work-study programs for approximately 1,200 National Youth Corps members in California colleges was not intended to be evaluatory; but he did report various indications of success, including administrators unanimously considering the programs successfful, participants having been predominantly positive in reactions, low dropout rates, and incomplete reports of C-plus grade-point averages during the following year in high school. Railsback (1970) presented a description of the Gary Job Corps Center program in San Marcos, Texas.

Warsh's follow-up study (1970) of behavior modifications in adults who had successfully completed a Michigan adult literacy training program between 1962 and 1966 revealed most could read at least as well as when they had completed the literacy-program, large numbers had continued to grow in reading and had been reading regularly, over 80 percent were involved in some form of organized activities (mainly education), and voting habits had changed.

Bryant, Bryant, Christiansen, and Fischer (1971) found a group of St. Paul basic education adults read significantly faster (115 versus 129 WPM) and with significantly higher (60 to 70 percent) comprehension (WRAT) when they read an SRA card while listening to a taped version than when they read another SRA card silently only. Cass (1970) discussed some main findings in a detailed analysis of teaching literacy skills to illiterate adults (below third-grade reading level) through, and proposed recommendations for more effective use of, Operation Alphabet.

Carpenter's study (1971) of adult literacy programs in Turkey revealed that negative self-image characterized program participants (but not nonparticipants); participation effected major attitude changes in participants; and participants became more active public media consumers, otbained new work experiences, and traveled more frequently and extensively. Opinions and attitudes otbained by Wali (1970) of Iraq learners who had participated in literacy programs using traditional verbal methods and those who had participated in alternative programs which included audio-visual techniques and vocational training revealed no differences in learners' choice of method, highly significant teacher preferences for audio-visual, and highly significant differences of learners' choices of literacy over vocational when presented individually.

Among findings obtained by Martinez (1971) in a questionnaire study of New York English programs for Puerto Ricans were that only 2 of 5 English programs used specific testing instruments, only 20% of teachers responding were native Spanish speakers, less than half of the teachers responding had taken a college course directly related to teaching English as a foreign or second language or to linguistics, and the audio-lingual method was reported as a preferred method by 57 to 75% of responding teachers.

In a study of effectiveness of remedial Algebra (not reading) programs in an Alabama junior college, Beck (1971) found no significant difference in course achievement between lecture discussion and multimethod groups; he did find the two groups significantly higher in course achievement than a programmed instruction group and large-group instruction as effective as small-group.

## FACTORS INFLUENCING READING AND OTHER STUDY HABITS AND SKILLS

### *Instructional Variables*

Among factors which appeared to Barrett (1970) to have promoted higher grades in first college English courses for University of Arkansas students were having been taught by teachers who were more highly qualified, with more favorable work loads, and having had appreciable writing experiences during the year prior to college English. Mallory (1970) found no significant differences in *Davis Reading Test* scores among groups for whom practice was provided under various conditions, involving two reading comprehension sub-skills (extracting and processing information) under two learning conditions (independent and supervised study). Pace (1971), after studying five teaching programs for teaching four severely retarded adults to read a set of four three-letter words, concluded that a "no-error rate" program was undesirable. Palmatier (1971) described a study in which the easiest of four note-taking procedures used with students was as effective for learning and achievement as were the other three methods.

Tobias and Abramson (1971) reported that, in a study involving programmed instruction, groups of City College of New York paid educational psychology students who used a reinforced constructed response mode on difficult technical material achieved more than did a group assigned to a reading condition with responses filled in, with anxiety effects noted on easy material but not on complex, technical content. Todd and Clemson (1971) studied the influences of meaningful material on recall by comparing the three response modes of reading (covert response), underlining, and note-taking (overt response) over three levels of difficulty (different lengths of a story) with University of Nebraska introductory psychology course students of high and low reading ability. They found consistent differences favoring better readers, and some differences favoring females, with regard to number of words, idea units, eight-word sequences, and identical words recalled, but few differences for difficulty level or response mode.

Test results otbained by Crewe (1969) with groups studying text material under differing review conditions indicated that retention could be increased by review and by studying close to examination time, indicating that cramming was effective when small amounts of study of material were involved. R. P. Williams (1970), after having University of New Mexico freshmen in reading and study courses select SRA Reading Lab cards and answer questions both before and after reading card selections, and finding a number of subjects could answer a considerable portion before reading selections, concluded that different amounts of comprehension were brought to the printed page by different subjects and that, unless severe comprehension blocks or problems existed, comprehension of subjects would increase from reading.

Mueller (1969) reported that an experimental group of 170 second-semester college French students, who had received both audio and visual confirmation of oral responses, scored significantly higher on writing tests than did a control group who had received only audio confirmation. Parent (1970) found Intermediate French students read

French selections with greater comprehension when they use parallel column, bilingual French-English reading material then when they use all-French reading materials; but use of a Language Laboratory, where students could hear French selections read, did not appear to enhance comprehension of French materials read. Findings obtained by Furbee (1971), in a study of writing skills of University of West Virginia freshman, indicated that drill and repetition types of instruction tended to result in less, rather than more, complex writing.

### *General Semantic/Grammatic/Syntactic Factors*

Begg (1971) concluded, after an experiment in which 16 volunteer female Canadian University students were asked to judge both meaning and wording of several hundred auditorially and visually presented sentences, that memory for meaning and memory for wording of concrete sentences were independent. Cohen (1970) found that when 12 paid volunteer graduate and undergraduate subjects cancelled visual targets (letters, phonemes, or words of a specific category) as they read prose passages semantic searching added little or no extra time when combined with visual (letter) or acoustic (phoneme) searching.

Forster (1970), in a series of experiments in which word sequences of varying syntactic arrangement and complexity were presented rapidly to paid volunteer undergraduate subjects, found that sentences containing only one underlying sentence were more accurately reported and concluded that perception was affected by sentence complexity only when the time available for sentence analysis was drastically limited. Foss (1970) found action time for noticing occurrences of a particular phoneme in a sentence longer for introductory psychology class undergraduate subjects when sentences were ambiguous, lexical ambiguities discovered more rapidly by students than structural ambiguities for groups required to notice both phonemes and ambiguities, and reaction time to monitor phonemes longer for ambiguous sentences when subjects themselves classified the ambiguity. When Johnson (1971) had University of Georgia Air Force ROTC students read passages of incremented difficulty and take cloze tests, he found a significant effect of passage difficulty on cloze scores, reading time, and cloze test time but little effect of punctuation on comprehension. Paivio's (1971) findings in studies in which Canadian University summer school students were exposed to English noun and object nominalizations (such as "growling lions", "raising flowers"), of high and low imagery, indicated positive effects on recall of imagery but not of deep-structured complexity (type of nominalization), with noun imagery alone being the best recall predictor. After having 18 volunteer undergraduate students read passages under one of six conditions of oral and silent reading and punctuation, and then read cloze tests over passages aloud, Weaver, Holmes and Reynolds (1970) found diffreent levels of difficulty producing different results on cloze and latency tasks but differing from earlier Bormuth findings. Reynolds and Lundrum (1970) reported a study involving step-wise analysis of word association data.

Groups of University of Texas introductory psychology course students were reported by Foss and Cairns (1970) to have learned complex sentences relatively less well than simple sentences as word lists also learned increased in length, with further indications that interruption of learning the word list was the important variable. When Kaplan and Cawellas (1971) presented introductory psychology students at Hofstra University with a sample of content ranging from 0 to 40 words, with one word deleted in each sample, and gave them 15 minutes to list all words which might be the deleted word, they found that as content increased the proportion of total words recallable increased, but the number of words which could fit the context decreased. Results otbained on a test with 60 sentences, given to 90 sophomores and juniors in a college course on reading instruction by Tuinman and Hafner (1971), indicated that position of adjectives (before and after noun) did not have significant effects on cloze test performance; but method of measurement (reading intact sentences before doing cloze test versus doing cloze test without pre-reading) did have a significant effect. Factor analysis of four grammatical and three word-length categories, which comprised the types of words deleted from a passage from which cloze scores of upperclass education students were obtained by Rentel and Kennedy (1970), found two different factors emerging, a long-word and a short-word factor. In a study of effect of sentence length and grammatical structure on repeatability of English sentences by native and foreign speakers, Stieglitz (1971) found influences of syntactic structure as well as of number of words, with substitutions comprising the greatest percentage of errors in eight-word, and omissions the greatest percentage in eleven-word, sentences.

Lawrence (1971a) found that subjects' (supposedly college students) estimates of words seen in a sequence of tachistoscopically presented words decreased with increased presentation rates, as did also number of words correctly recalled. Lawrence also found (1971b) that when subjects searched word lists for target words, under conditions of various rates of tachistoscopic presentation, such words were readily identified at all rates when the target word was the last one in a list; but when it was the first one, or imbedded in the list, accuracy of recognition was an inverse function of presentation rate. Subjects also did significantly better when target words were identified with a category then when non-categorical. McEwen and Greenberg (1970) found that when groups of Michigan State University students in an introductory communications course read either "high" or "low" intensity versions of a message on teethbrushing practice, the high intensity messages tended to be regarded as clearer and more logical and sources to be more trustworthy, qualified, and dynamic.

In a series of investigations of retroactive interference with prose-learning tasks (200-plus-word passages), in which changes of segments of original passages were interpolated, Crouse (1971) considered a major finding to be that retroactive inhibition was influenced by inter-passage similarity and by more than passage differentiation. Paivio and O'Neill (1970), in a series of experiments involving tachistoscopic presentation of stimulus nouns to ten Canadian undergraduate students in introductory classes, found frequency (Thorndike and Lorge) and familiarity had strongest influence on visual recognition ease while the image-evoking characteristic was unrelated. May and Tryk (1970) found rare words were recalled more frequently than common words when a mixed list of alternately common and rare words was presented, which was the reverse of the typical high-frequency-easier-recall effect. Loewenthal (1971), in word association experiments with University of North Wales undergraduates, involving twenty seven-letter nouns from the Thorndike-Lorge word list, found equally ready recognition of "known" and "familiar" words which were tachistoscopically presented, with known words eliciting semantically related, and "familiar" words phonologically related or mediated, responses; similar, but to a lesser degree, findings were obtained for "unknown" words.

### *Personal/Personality Factors*

Remedial English students in Beyer and Oetting's experimental group (1970), who were given initial encouragement and reassurance about their potential achievement, did not differ significantly from a control group in grades in smaller classes; but in larger classes their grades were significantly lower, leading to the conjecture that positive encouragement effected lower anxiety, which led to poorer performance. Stable introverts (Eysenck Personality Inventory) among groups of British technical college students were found by Cowell and Entwistle (1971) to have better study habits (Brown-Holtzman) but not significantly better examination results, than a group of extroverts. Various SSHA subscale scores were significantly related to academic achievement, but in contrasting personality pattern relationships. Defrain (1971) found that none of the California Test of Personality subscale scores of students who completed an Oklahoma State University reading program were significantly different from those who had dropped out of the program. Englander "self-concept as a reader" and "self-concept as a student' 'scores were significantly different for two male and two female groups, respectively.

In a study undertaken by Jamieson (1970) with ten Catholic and ten non-Catholic New Zealand teachers college students, who were presented with controversial religious material, the hypothesis that possible pro-statements and implausible anti-statements would be learned more easily than opposite types was confirmed for pro-Catholic, but not for anti-Catholic, subjects. Analysis of various scores obtained on Wayne State University marginal admission students ,by Pepper and Drexler (1971), indicated that students who felt an "internal locus of control" tended to report better study habits and attitudes than those who felt an external locus, showed negative relationship betwen locus of control scores and SSHA Scores, and revealed an interaction between locus of control and instructional method used in a reading program. W. A. Smith (1971) found no relationship between Personality Assessment Systems scores of University of Missouri rapid reading course participants and enrollment, attrition, or other characteristics; but he did find indications of relatve influence of teaching styles of instructors.

When words representing the subjects' highest and lowest value areas (Allport-Vernon-Lindzey) were presented stereoscopically to fifteen Brooklyn college introductory psychology class students, Gaertner and Seidenberg (1971) found that students generally recognized words of idiosyncratic familiarity. Gelso, Osterhause, and Bodden (1971), studying results obtained after administering various scales to University of Maryland students in an educational skills course and grade-point improvement, concluded that the extent to which students profit from skills improvement courses is dependent on the student's pre-course level of performance. Weigel (1971), using a laterality test and devised mirror-writing and mirror-reading tests with 22 college students, concluded that the three characteristics were related, with the strongest relationship appearing to be between ambilaterality and mirror-writing.

In a study of the impact of editorial satire directed against a person (Art Buchwald editorial columns regarding President Nixon and Vice-President Agnew), carried on with 156 University of Georgia beginning speech course students, Gruner (1971) found negligible effects on character and authoritativeness ratings and by humorousness/lightness ratings. When experimental groups of Illinois State University students in a beginning speech course were exposed to either a reward-oriented or punishment-oriented message (emphasizing benefits or losses derived in education, medical research, air and water pollution control or space program appropriations), McCroskey and Wright (1971) did not find differential effects of type of approach on either attitude change or perceived credibility; but they did find considerable attitude change effected by high-credible, but not low-credible, sources.

*Media/Graphic/Visual Factors*

When Anderson (1969) presented each of four concepts to approximately 1,600 undergraduate speech class students in various groups in pictorial, written, or spoken media, or one of the various combinations, he found that single-medium presentation elicited greater strong connotative associations than did multi-media presentation; and he concluded that media which use the same sign types competed with one another when combining in multi-media presentations. In a Dwyer study (1971c), involving 261 college students and four different types of black and white and colored illustrations accompanying a taped presentation on the human heart, the different types of accompanying visuals failed to facilitate terminology and comprehension test achievement; but performance on a drawing test and on a total criterial test were favorably affected. However, the types of visuals students judged being most effective for them were not always the ones actually found to be the most effective. In a study involving presentaiton of a reading test in four different type faces, and multiple-choice item options presented either "across" or "down", evidence obtained by Follman, Lowe, and Miller (1970) indicated that neither type face nor item options significantly affected the test performance. In another study (which was related to but not directly involved with, reading) Follman, Lowe and Miller (1971) found that type face did not significantly affect grades awarded to themes of upperclass educational psychology students.

*General Readability and Retention Factors*

In an "empirical, quasi-experimental" investigation of effects on a lay audience of textual variables in science articles written by a number of different writers, Funkhouser and Maccoby (1971) found strong influences of readability and use of example but did not find the typical differences between sexes in reading enjoyment or between science majors and non-majors in understanding of content; considerable variation was found among professional science writers of the effects of their writings on lay audiences. Newsome and Gaite (1971) found no significant differences in immediate retention between groups of students who had read a 2300-word science fiction selection and a group who had read a 300-word passage based on information retained over one week by subjects who had read a longer selection; but delayed retention (one week) was significantly better for shorter passages than for longer ones. Hardyck and Pertrinovich (1970), in a study carried on with 18 freshmen remedial English students who had muscle actions while reading recorded (electromyograms) and who were required (according to group placement) to keep laryngeal or forearm flexor muscle action relaxed, found the laryngeal feedback group doing less well on comprehension of difficult material than did the other groups. Weaver, Kingston, Bickley, and White (1969) found best predictors of reading comprehension (*Davis Reading Test*) and post-cloze scores of 53 psychology class sophomores to be *Modern Language Association Test* subtests which required subjects to learn previously unfamiliar material by rote. Smith and McCombs (1971) found that, while improved reading ease (Flesch) increased preference for a news story, increased white space appeared to make for more differences in appeal for story; but comprehension of content was not affected by different versions.

### *Mathemagenics*

Performance of undergraduate educational psychology students in Frase and Silbiger's study (1970), in which four experimental groups were asked to search passages for specific information, indicated that as subjects search for information they establish criteria for discerning relationships among separate sentences, both for sentences which identify specific information and for those which must be used to reject relevant information. MacGregor (1971) concluded, after a study with "highly self-selective" adult subjects, that position of questions in a reading selection (before, after, or interspersed) was not a significant factor in long retention of learning (six to eight days) and that general reading ability was a more significant variable.

In further reports of Dwyer's studies of effectiveness, on learning of undergraduate college students, of visuals accompanying textual presentation concerning the heart Dwyer (1971b) indicated that questions as advanced organizers in visual instruction (questions preceding each visualized page in instructional booklets) were not effective instructional techniques for achieving various types of objectives involved. In still another report, Dwyer (1971a) indicated that when questions followed textual material presented, visualizaton of content material was not effective as an instructional technique.

In a study concerned with differences and quantity of output (ideational fluency) produced by various modes of input, and involving 140 adult basic education students, Reynolds, Palmentier, and Ulmer (1969) found fluency (writing responses associated with given words) not a function of presentation modes (auditory, visual, graphic, or combinations).

Baldwin and Bailey (1971), in a study to determine influence of a microfilm presentation method on information processing (visual modality) ability, prepared three types of tests for each of four types of Air Force training material and found that Air Force subjects receiving offset copy materials consistently scored significantly higher than did subjects receiving materials via positive image or negative image microfische.

Weigel, Weigel, and Herbert (1971), carrying on an investigation in line with the suggestion that generalizations made in research projects where students were required to participate (as a course requirement) may be legitimate but may not be warranted, let 100 introductory psychology students (from a pool of 1000) sign up for study during the first, middle, and last week of the term. They found differences in students study habits (Brown-Holtzman) which they felt could affect generalizability of research findings.

### *Factors Influencing Listening*

Responses obtained by Bacharach and Kellas (1971) from undergraduate general psychology students fulfilling a course requirement, who were asked to recall, in writing, each of a number of auditorially presented sentences, indicated that recall was a function of phrase structure of the sentence material rather than a function of types of sentences used (active, passive, and manner adverbial). Carey (1971) reported that when students and nurses listened to stories recorded at one, two, and three words per second, shadowers (those who repeated auditory messages aloud as they heard them) had somewhat higher word recognition and semantic retention scores at the slowest rate than did listeners; but at faster rates the difference disappeared. Word recognition scores for shadowers were found related to either semantic retention or syntax recognition. Hakes and Foss (1970) found that when 40 introductory psychology students heard self-embedded sentences they had faster phoneme monitoring time (listening for a word beginning with a given letter) when relative pronouns were present in sentences than when these were deleted. They concluded that a phoneme monitoring technique is a better index of comprehension difficulty than is a paraphrasing technique. Data obtained by Hamilton and Deese (1971) when 30 subjects listened to 288 tape-recorded sentences and rated them as to comprehensibilty indicated that grammatical structure, rather than semantic relationship, is a chief determiner of comprehensibility. Cairns and Foss (1971) found no effective word frequency for nouns or verbs when undergraduate psychology students (required to participate in an experiment as a course requirement) pushed a button when they heard words beginning with a given phoneme among 45 auditorially presented sentences (but some effect was found for low frequency adjectives).

Results obtained by Rossiter (1971) with groups of undergraduate introductory communications course students, who listened to various types of recorded information played at varying speeds, showed interaction effects between speed of presentation and type of information, with significant declines in comprehension of compressed speech at rates below 275 words per minute. In another study, Rossiter (1970) obtained a correlation of .46 between age and listening scores of 30 adults, leading him to conclude that older students might retain less from oral presentations than do younger students.

Sticht (1971) in another of his experiments with time-compressed speech recordings, done with 75 high- and 75 low-aptitude men (*Armed Forces Qualification Test*) found that new material was learned in the time saved by the compression of speech, but not more than that which resulted from listening to uncompressed material. Watts (1971) found no differences between lecture test results for U. S. Air Force groups which had heard lectures in a normal auditorium situation and groups which had heard versions compressed to 250 words per minute; but he did find significant effects relative to rank and to subject matter.

Warren and Obusek (1971) found phonemic restoration (the illusory effect which occurs when a missing speech sound in a sentence is restored by the listener on the basis of prior and subsequent conduct) occurring much less frequently for 100 introductory course Wisconsin students on a silent gap rather than an extraneous sound (such as a cough or tone) replace the missing speech sound. Results obtained by Guthrie (1971) when 90 undergraduate, male, paid subjects were given several types of feedback while learning 39 whole sentences verbatim indicated feedback facilitated learning when it followed wrong, but not when it followed correct, responses; but feedback consisting of both stimulus and response was superior to no feedback; and there were no differences between only-response feedback and no-feedback conditions. In a study involving 26 undergraduate college students (presumably in a psychology class) taking notes on a 160-word taped passage, Howe (1970) found that recall success was considerably dependent on the extent to which recall items were reproduced in notes, with the chances of subsequent recall of items not appearing in notes being very low. Meyer and Miller (1970), in a comparison of listener information retention and attitude evaluation to emphasis (vocal stress and intention of pauses to set off key words and phrases) and non-emphasis versions of a radio newscast, did not find a significant effect of delivery style on recall, although attitude measures tended to show a preference for emphasis delivery.

## INFLUENCE OF READING AND STUDY HABITS AND SKILLS

Reactions of subjects in a program carried on with chronic-emotionally-disturbed patients in two Washington state mental hospitals led McClaskey (1971) to conclude that bibliotherapy could be effective in bringing about significant behavioral changes but not significant attitudinal change gains. Experimental evidence obtained by Atkin (1969) with groups of University of Wisconsin undergraduates indicated te effectiveness of exposing voters to opinion poll reports which showed candidates or issues highly favored by a specific electoral population. Questionnaire responses obtained by Brahee (1970), from over 500 adults enrolled in a course at the University of Michigan Adult Extension Center, revealed that respondents relied more extensively on newspapers and magazines for meeting information needs than on television and radio, with further indication that higher as well as lesser educated enrollees used television primarily for entertainment. Results obtained in a study by Hiett (1969) of the effectiveness of gun control advertising, carried on with Southern Illinois University sophomore sociology students, led Hiett to conclude that attitudes toward gun control legislation were little influenced by single exposure to pro-gun control advertisements but that knowledge of gun ownership, size of hometown, and political leaning would be better predictors of attitude.

In a study of 104 male adults in a Candian retraining, or academic upgrading, program, Sainty (1971) found that slightly over half who dropped out had lower intelligence and lower *Gates Reading Survey* scores than the completers, but not lower personality scores. Vraa (1971) found ACT Science Reading and Natural Science Reading scores to be significant variables in the academic success prediction of 101 Canadian male freshmen at the University of North Dakota; the best individual predictor was ACT English scores rather than high school grade-point-average, which was found to be the best predictor in other studies of American students.

Leslie and Calfee (1971) found, in a study carried on with 8 college students from introductory psychology courses as well as 48 elementary children from remedial reading and mathematics classes, that reading ability was not a significant factor in various visual search tasks involving scanning lists of words. Ross (1971) found Cooperative English Reading Speed Comprehension scores not useful in predicting first-semester achievement of University of Akron engineering freshmen. Carbuhn (1970) found reading speed (*Gates Reading Survey*) least significant and overall reading average correlating most highly with aptitude and GED scores of 157 Utah Urban Job Corpsman; and a reading test score of approximately Grade 7 level was found necessary for passing various GED tests. Analysis of test data by Hansome and Taylor (1970), including ACT Reading Test scores obtained for over 400 Minnesota Technology Institute freshmen engineering students, indicated that, while ability (including ACT reading scores) will influence academic success, personality characteristics (Minnesota Counseling Inventory) will influence students withdrawing or remaining in college.

## READING AND RELATED HABITS, TRAITS, AND SKILLS

### *Study Techniques and Habits*

Adams (1970) did not find significant differences in the amount of underlining done by students in undergraduate methods textbooks or among amounts done in texts for different areas. Positive, but relatively low, correlations were found between amount of underlining and grades in specific courses, but lower correlations for reading than for social science and science courses. When total amount of study time and time between original study and review was controlled, Hultgren and Crewe (1969) found no significant differences in retention of textbook study material between University of Minnesota student athletes who crammed and those who employed study-review strategies (including underlining). Maxwell (1970 found that advanced students with average reading skills who participated in scanning practice and study were able to improve scanning speeds, develop different strategies for learning the scanning task, and did not show significant differences between scanning speeds and stimulus presentation method (auditory versus visual). Wark (1971b) described a study in which students in a how-to-study course, given a conditioning treatment for relaxation and desensitization for reducing test anxiety, dropped an average of 20 points on an anxiety scale.

Among findings obtained by Sommer (1969) in his interview-visit survey of use of students' rooms for study activities (conducted on eleven university, college and junior college campuses) were that a person with a roommate was more likely to study if the roommate were physically present and studying than when the roommate was out of the room; fewer students in apartments studied at their desks in double and single rooms (40% at desks, 40% on beds—of whom only one half were studying, and remainder on floor or at other locations); no gradepoint differences between those who studied at desks and those who studied on beds; and that the most important study aspect of residence hall rooms for students was the roommate, not the furniture. Santiago (1971) found that Spanish speaking students in Puerto Rican colleges wrote in similar fashion (relative to numbers of paragraphs, sentences, and words) in Spanish as in English essays. Yarington (1970) presented a review of 56 studies, of approximately the last 90 years, of time spent reading and studying by college as well as high school students.

### *Reading and Verbal Ability Levels and Habits*

According to the *Phi Delta Kappan* (1970), survey figures gathered at the request of the National Reading Council indicated that 7% of Americans have difficulty with filling out a Social Security application form, 11% with a personal bank loan form, and 34% with a Medicaid application. Over 4 million Americans are considered functionally illiterate in present society, with functional illiteracy ranging from 4 to 16% for rural and from 2 to 9% for suburban residents. Highest region of illiteracy was in the South (4 to 15%) and next in the East (4 to 14%), with 3 to 13% in Midwestern areas, with fewer reading deficiencies alledgedly found in the West than in other regions. D. A. Brown's (1970) analysis of data, obtained by administering the WAIS to 128 functionally illiterate adults (of whom over 95% were black) enrolled in adult basic education courses in city core areas, indicated that, while the tested groups scored more poorly on the WAIS than literate whites would be expected to, the black disadvantaged group did not show comparatively greater strength on rote learning than on abstract thinking tasks. When Knight and Alcorn (1970) compared the intelligence scores and reading performance scores of 24 adult education program enrollees with elementary subjects roughly matched on the basis of California Reading Test scores, they found stronger relationships between intelligence (IPAT) and reading performance scores for elementary than for adult subjects, with adults performing as well or better on one reading test than on a cloze test and considerably better on another test than elementary subjects. They interpreted results to suggest that adults might be able to utilize reading skills they do possess to a greater extent than standardized test results might indicate.

Dinnan, Bickley, and Cowart (1971) found college students with poor verbal ability (and considered "poor risk") responded with higher syntagmatic responses (all verbal responses which are not "superordinate, coordinate, whole-part or area contrast") on the *C/A Oral Language Inventory* than did students with high verbal ability. Gibson (1971) reported that *Nelson-Denny Reading Test* scores of a sample of 200 California community college students indicated that 65% had poor reading ability, 24% average, and 11% superior, with informal reading inventory results indicating that over half were reading at frustration level. Readability levels of the students' textbooks were found to range from just below mean reading level to approximately four-and-one-half years above the students' mean reading levels. When McClellan (1971) compared reading levels (Nelson-Denny) of Hillsborough Junior College students in four social science area classes and one English class with readability levels (Dale-Chall) of required texts, she found the readability levels significantly higher, with analysis of readability levels of twenty selected texts indicating that less than 30 per cent of students enrolled would be able to read their texts. Questionnaire data otbained by Williams (1970), from over eight-hundred Harrisburg Community College freshmen, revealed only a moderate correlation (.40) between total books read and English and social studies scores (ACT), with over half reporting having read fewer than fifteen to twenty books before college entrance. Half the non-readers reported having no reading difficulty, while about a third of the readers reported difficulty. Electromyographic chin-lip-sight recordings, obtained by Locke (1971), of 6 young adults as they silently read three different passages containing high, medium, or low proportions of words whose pronunciation required labial motor gestures (*p, b, m, w, f, v*) indicated significantly higher tracings for high- than for low-labial passages, suggesting that covert oral activity during silent reading is a measurable form of speech.

### *Use of Reading and other Media*

Personal interview data obtained by Block (1970) from three-hundred-fifty St. Louis inner-city-poor subjects revealed that mass media do reach the poor, particularly television and newspapers, with over 70% of respondents indicating they viewed TV regularly and slightly less claiming they read or at least looked at a newspaper almost daily. Kline's (1971) use of a path analysis technique, in efforts to find the extent of influence of life style and demographic factors on extent of media use, revealed TV taking up 29 hours per week of time spent by a sample of urban and suburban mothers and fathers, radio 25 hours per week, newspapers 11 hours, magazines and books 7 and 6 hours, respectively, with education correlating negatively with newspaper reading. Among questionnaire findings obtained by Dulin and Otto (1970) from evening reading improvement course enrollees were that "fiction' 'and "technical" were types of books most often read; a daily average of approximately 32 minutes was reported as being spent reading newspapers; meeting leisure reading needs, reading needs of current jobs, and reading needs for future intended schooling were cited as major reasons for desiring to improve reading skills, with nearly all feeling that reading speeds could be increased over four-hundred words per minute by average readers. Rogers (1970) found "leaders" among a group of former Soviet residents indicating they would read a higher proportion of editorials, analyses, and letters in leading Russian newspapers (*Pravda* and *Izretia*) than did "nonleaders", similar to Western Soviet scholars.

### *Teacher and Other Adult Strategies and Attitudes*

Browning (1971) found no significant differences in increase of knowledge of phonics generalizations among groups of college students enrolled in courses in reading and language arts who were exposed to one of three two-week treatments (individualized audio-visual, combination of individualized audio-visual and lecture, and lecture only). Illika (1970, 1971) has presented results of another of his surveys of phonic skills of teachers and reported finding no significant differences among groups of graduate students who were teachers, undergraduate elementary educations students, and undergraduate speech and speech therapy students. Samuels and Chen (1971), exposing 25 elementary children and 25 college students to various word pairs and speed of word recognition and word recognition strategies tests, found adults superior with every strategy used.

Smith, Hesse, and Otto (1970), in a continuation of earlier reported work on an inventory of teachers attitudes, found junior high English teachers expressing most positive, and science teachers least positive, attitudes toward teaching reading in their content area classes. Questionnaire responses obtained by Otto and Dulin (1970) from approximately one-third of the 237 National Reading Conference members polled, concerning status-ranking of various professional activities, indicated relatively little differences between "large university" and "others", "research" and "no research", and "teacher preparation" and "reading service" groups, with journal research reports, writing professional books, and leadership in national level organizations tending to carry major status.

The answer to what college students are reading currently is (according to *Parade*) "comic books", with national best sellers including "The Mighty Avengers", "The Incredible Hulk", "The Fantastic Four", and "The Amazing Spiderman" (Swift, 1971). While still featuring super-heroes whose survival is accomplished by supernatural intervention, the comic books, reportedly, "have been transformed from the purely 'escape literature' of Flash Gordon to a relevant new and interesting art form", "even social commentary", and "have become so papular that they run the risk of acceptance by the Establishment [Swift, 1971, p. 10]."

## MISCELLANEOUS

### *Tests and Measuring Procedures*

Carver and Darby (1971) described the revision and construction of two forms of the chunked type of tests developed and analyzed earlier by Carver (designed to show informatiaon stored during reading). They reported that administration of the test to groups of paid college students yielded results which indicated that the test was a valid measure of differences in information stored between readers and nonreaders of passages. Farr and Smith (1970) discussed a procedure for assessing test items and the effect of specific items on test validity and reliability. Results obtained did not sustain the hypothesis that dropping four items from the *Nelson-Denny Reading Test* would increase test reliability and validity. Grotelueschen and McQuarsie (1970) concluded, on the basis of correlations obtained between scores of 178 adults on the *Quick Word Test* and WAIS scores obtained two years earlier, that the QWT was a valid test for measuring adult mental ability. A further report of Follman's analysis of three critical thinking tests, reported at an earlier conference, was also presented (Follman, Miller, and Burg, 1971).

Rankin and Hess (1970) presented a method for determining internal (intraarticle) reading flexibity which involved obtaining relative reading rates along with determining relative difficulty of successive small portions of passages. Trial of the procedure yielded "flexibility quotients" which indicated relative inflexibility and which gave some reason to question validity of sub-group flexibility coefficents. Taschow (1971) has presented results of further study of reading gains, those measured by the conventional crude gain method versus those measured by the residual gain method, which indicated that crude gain measurement tends to underestimate as well as overestimate individual differences in reading improvement and that residual gain procedures tend to eliminate overestimates of the inferior improvers and underestimates of the superior improvers. Using college introductory communications course students as judges, Sachsman (1970) explored a method for analysis of "loading" with regard to referents in headlines, stories, and combinations. All six Bay area newspapers involved were judged as showing significant bias, with "pro" loadings for President Nixon as a referent and "con" loadings towards Arab Commandos, offshore oil companies, University of California (Berkeley) demonstrators, and considerable lack of agreement among judges with regard to striking teachers and state college trustees. McLeod and Anderson (1970) described a series of experiments, which utilized both adults and children. and mathematical derivations which lead to an "uncertainty reduction index" (URI), which purportedly could be regarded as a valid and reliable expression of reading ability of subjects. Askov (1970) discussed the revision of a survey of teacher attitudes toward individualized reading instruction and suggested the instrument might be useful despite extraneous influence of survey administrators or uncertain interpretation of scores. Blake (1971), in a study conducted with undergraduate history students and cloze procedures, found agreement scores (ratio of amount of agreement among responses to a completion item to the greatest amount of agreement possible) more sensitive to differences in contextual strength than were cloze scores.

In a series of studies involving measurement of pupil dilation of college subjects while reading or listening to selections of various difficulty levels, Carver (1971) found mean reading time increased with difficulty; but he did not find that pupils of students dilated while reading or while listening to aurally presented connected discourse. Kanungo and Dutta (1971) have described their perceptual framework for communicative meaning of words and their development of the Semantic Consensus Index (SCI), a technique for measuring relationships of communicative meaning of words with such other meaning attributes as imagery concreteness, meaningfulness, dictionary meaning, and intensity of affective meaning. Lynch (1970) reported evaluation study of an Index of Syntactic Dispersion (ISD), a method of stylistic analysis or "a measure of the variance in strings of three parts of speech." Mickleson (1971) presented a table of "meaningfulness," based on data from groups of fourteen-year-old Canadian children and Canadian adults, for 120 nouns.

*Magazines and Newspapers*

Application of the Flesch readability formula to lead articles in issues of leading popular magazines (*Reader's Digest, Saturday Evening Post, Ladies Home Journal, Harper's,* and *Popular Mechanics*) at five-year intervals, from 1925 to 1965, by Barganz and Dublin (1970) revealed that the articles, from earliest issues on, required at least high school reading level ability, and over half of the articles checked required greater than that. Mujahid's content analysis of coverage of Pakistan in three leading U.S. news magazines (1970) showed the types of stories to have been primarily ones concerned with Pakistan's foreign relations (with meager coverage of domestic matters) and that the reportage in the magazines was, in general, factual, balanced, and objective. Stempel (1971) found significantly less percentage of news pictures with only whites in 1970 issues of five leading U.S. news-picture magazines than in 1960 issues and a significantly greater percentage of whites-only pictures for advertising than for news pictures. Geizer's comparison (1971) of advertisements in 1960 issues of *Ebony* and *Life* with those in 1969 issues indicated significantly lower frequency of light-skinned-black model and of white-only model advertisements, significantly greater frequency of dark-skinned-black model and of mixed model advertisements, and no significant changes in use of "black" rather than "Negro" or in "black identity" language, leading to the conclusion that drastic identification with black militancy had been rejected by Ebony, a magazine regarded as a leading black mass circulation magazine. Ads with blacks in twelve general circulation magazines studied by Kassarjian (1969) were found to be three times greater in 1969 than in 1964 and in 1965, but blacks still were found in less than 2 percent of the ads.

M. R. Brown (1970) concluded, on the basis of a study in Chile of the effect of mass media content on Chilean farmers' behavior, that "illiteracy, lack of education, and so-called traditional attitude did not constitute impenetrable variables to the flow of technical information [p. 734]." When Cote (1970) compared 37 pairs of stories, he found no differences in accuracy between AP-originated and UPI-originated stories; and he concluded that errors found were due more to reporting mistake sources than to interpretation errors. Exhaustive theme content analysis of editorials and letters in the Bridgeport, Connecticut, *Post* by Spector (1970) indicated that the newspaper had a definite influence on referendum voting and lead to the conclusion that "a newspaper with a reputation for forthright editorializing on local politics will probably be more influential on its readers than those papers which avoid local controversy [p. 766]." Little (1971) has presented figures, based on newsprint consumption, to show that the growth of city newspapers has been in inverse relation to size of populations, ranging from a 41.3% increase in cities of one million or more to 93.2% in cities of less than 100,000.

### *Materials*

Bahe (1970) presented an analysis of college reading manuals or practice materials and a cross-reference listing of materials for various specific reading or study skills areas. Geeslin (1970) field-tested multilevel reading materials with adult basic education classes. Teachers, teacher-trainees, and adult education experts reportedly rated the materials as just as adequate or definitely more adequate than traditional materials.

### *Eye Movement Studies*

Geyer (1970) briefly reported (with detailed evidence and information presented elesewhere) a method for obtaining a coefficient of constancy of the temporal eye-voice span as a measure of overall reading effectiveness. In an experiment involving detailed measurement of eye movements of three subjects while reading three passages of varying difficulty, Smith, Schremeser, and Putz (1971) obtained evidence that the eyes are not "perfectly conjugate" in saccadic motion, that the left eye tends to lead the right eye, and that time difference was not related to difficulty of materials but that change in horizonal rotation of reading display was. Singer (1970) found, when groups of New Mexico State University students were tested with the *Cooperative English Test* under one of three conditions, that when reading was paired with listening, subjects made more regressions, had smaller recognition spans and longer durations of fixations, and reading rate was decreased (as it was also with silent reading).

### *Other*

Meehan (1971), studying groups of senior language arts methods students who were given instruction on question fluency by traditional approaches, by critical reading with emphasis on affetcive dimensions, or by critical reading with emphasis on cognitive dimensons, found no significant differences among the groups with regard to cognitive-memory or critical-type questions used in teaching groups of children. Tichenor, Olien, Harrison, and Donohue (1970) found communication accuracy (as judged by respondents immediate recall statements of what they had read) for 73 science articles pulled from midwest metropolitan newspapers to range from 30% to 91%, with an average of 65%, with accuracy being higher for editor- assigned articles than for those originating with public meetings or on a reporter's own initiative. Mustico (1971) presented some implications for reading readiness on the basis of responses of college sophomores in educational psychology classes and students in Grades 3, 6, 9, and 12 to sets of three-letter trigrams. Of 39 variables studied by Yang (1971) in developing a readability formula for the Chinese language, highest predictive power was found for, in order, proportion of "simple" (in list of 5600) words, average number of strokes per character, and proportion of complete sentences.

## REFERENCES

Adams, E. K. Underlining: the graphical aid to college reading. Schick & May (1970a, 12-22).

Ahrendt, K. M. An adult reading improvement program. *Reading Improvement,* Spring 1971, *8,* 7-10.

Askov, E. N. Assessing teachers' attitudes toward individualizing reading instruction. In Schick & May (1970a, 23-28).

Atkin, C. K. The impact of political poll reports on candidate and issue preferences. *Journalism Quarterly,* Autumn 1969, *46,* 515-521.

Anderson, J. A. Single-channel and multi-channel messages: a comparison of connotative meaning. *Audio-Visual Communication Review,* Winter 1969, *17,* 428-434.

Bacharach, V. R., & Kellas, G. Phrase versus base structure effects on short-term retention. *Journal of Verbal Learning and Verbal Behavior,* April 1971, *10,* 171-175.

Bahe, V. R. A content analysis of current college reading manuals. In Schick & May (1970b, 11-25).

Baldwin, T. S., & Bailey, L. J. Readability of technical training materials presented on microfische versus offset copy. *Journal of Applied Psychology,* Feb. 1971, *55,* 37-41.

Barganz, J. C., & Dulin, K. L. Readability levels of selected mass magazines from 1925 to 1965. In Schick & May (1970b, 26-30).

Barrett, C. J. The influence of high school senior english teachers and the senior courses in high school english on students' performance in freshman composition at the University of Arkansas. (Doctoral dissertation, University of Arkansas, 1970.) *Dissertation Abstracts International,* Sept. 1970, *31,* 919A. Ann Arbor: University Microfilms, Order No. 70-17, 173.

Beck, M. C. A comparative analysis of three methods of teaching remedial algebra on the junior college level. (Doctoral dissertation, Auburn University, 1970.) *Dissertation Abstracts International,* June 1971, *31,* 6270A. Ann Arbor: University Microfilms, Order No. 71-13, 196.

Begg, I. Recognition memory for sentence meaning and wording. *Journal of Verbal Learning and Verbal Behavior,* April 1971, *10,* 176-181.

Berger, A. A comparative study of reading improvement programs in industry and education int he United States and Canada. In Schick & May (1970a, 29-33). Also in Wark (1917a, 14-37).

Beyer, H. N., & Oetting, E. R. Effect of encouragement and reassurance on students' performance in remedial english. *Journal of Educational Research,* Oct. 1970, *64,* 59-60.

Blake, J. H. Cloze scores, agreement scores, and approximate equivalence scores as indices of contextual constraint. (Doctoral dissertation, Syracuse University, 1970.) *Dissertation Abstracts International,* May 1971, *31,* 5923A. Ann Arbor: University Microfilms, Order No. 71-10, 892.

Blevins, M. L. A comparative study of three methods of instruction in vocabulary on achievement of students in the adult institute. (Doctoral dissertation, Oklahoma State University, 1970.) *Dissertation Abstracts International,* April 1971, *31,* 5108A. Ann Arbor: University Microfilms, Order No. 71-11, 105.

Bliesmer, E. P. 1969 review of research on college-adult reading. In Schick and May (1970a, 46-78).

Bliesmer, E. P. 1970 review of research on college-adult reading. In Greene (1971, 361-389).

Block, C. E. Communicating with the urban poor: an exploratory inquiry. *Journalism Quarterly,* Spring 1970, *47,* 3-11.

Brahee, C. I. A study of mass communications influence upon the adult learner—implications for programming. (Doctoral dissertation, University of Michigan, 1969). *Dissertation Abstracts International,* Aug. 1970, *31*, 601A-602A. Ann Arbor University Microfilms, Order No. 70-14, 482.

Brooks, R. B. Remedial education in Surrey. *Adult Education* (London), Sept. 1970, *43*, 150-156.

Brown, D. A. Intelligence of Adult Illiterates. In Schick and May (1970a, 94-98).

Brown, M. R. Communication and agricultural development. *Journalism Quarterly*, Winter 1970, *47*, 725-734.

Brown, W. F., Wehe, N. O., Zeinker, V. G., & Haslam, W. L. Effectiveness of student-to-student counseling on the academic adjustment of potential college dropouts. *Journal of Educational Psychology,* Aug. 1971, *62*, 285-289.

Browning, M. E. P. An analysis of different learning strategies for the achievement of phonics generalizations. (Doctoral dissertation, Auburn Univesrity, 1970). *Dissertation Abstracts International,* April 1971, *31*, 5242A. Ann Arbor: University Microfilms, Order No. 71-9036.

Bryant, A. S., Bryant, B. F., Christiansen, D., & Fischer, L. The effectiveness of taped reading instruction in increasing the reading rate and comprehension of undereducated adults. *Adult Education,* Summer 1971, *21*, 246-252.

Cairns, H. S., & Foss, D. J. Falsification of the hypothesis that word frequency is a unified variable in sentence processing. *Journal of Verbal Learning and Verbal Behavior,* Feb. 1971, *10*, 41-43.

Caple, R. B. Group study of low-achieving freshman males in a residence hall setting. *Journal of College Student Personnel,* May 1969, *10*.

Carbuhn, W. M. Predicting general educational, development (GED) test performance of urban job corpsmen. *Educational and Psychological Measurement*, Winter 1970, *30*, 993-998.

Carey, P. W. Verbal retention after shadowing and after listening. *Perception and Psychophysics,* Feb. 1971, *9*, 79-83.

Carpenter, H. F. Adult education and the transformation of rural society in Turkey. (Doctoral dissertation, Stanford University, 1970.) *Dissertation Abstracts International,* May 1971, *31*, 5925A-5926A. Ann Arbor: University Microfilms, Order No. 71-12, 871.

Carson, J. E. The effects of programmed instruction as a supplementary teaching aid in adult basic education in the Ohio State Reformatory. (Doctoral dissertation, Ohio State University, 1970.) *Dissertation Abstracts International,* Mar. 1971, *31*, 4441A. Ann Arbor: University Microfilms, Order No. 71-6951.

Cartwright, H. D. Study skills for the severely retarded college reader. In Schick and May (1970b, 50-57).

Carver, R. P. Pupil dilation and its relationship to information processing during reading and listening. *Journal of Applied Psychology*, Apr. 1971, *55*, 126-134.

Carver, R. P., & Darby, C. A. Jr. Development and evaluation of a test of information storage during reading. *Journal of Educational Measurement*, Spring 1971, *8*, 33-44.

Cass, A. W. The role of television in reaching illiterate adults with a literacy program series. (Doctoral dissertation, Columbia University, 1969.) *Dissertation Abstracts International,* Aug. 1970, *31*, 602A. Ann Arbor: University Microfilms, Order No. 70-12, 544.

Christ, F. L. A pre-college summer reading/study skills lab. In Schick and May (1970b, 58-64).

Christ, F. L. (Ed.), *Interdisciplinary aspects of reading instruction*. Western College Reading Association 4th Yearbook. Los Angeles: the Association, 1971.

Cohen, G. Search times for combinations of visual, phonemic, and semantic targets in reading prose. *Perception and Psychophysics,* Nov. 1970, *8*, 370-372.

Cote, J. R. A study of accuracy of two wire services. *Journalism Quarterly,* Winter 1970, *47*, 661-666.

Cowell, M. D., & Entwistle, N. J. The relationships between personality, study attitudes, and academic performance in a technical college. *British Journal of Educational Psychology*, Feb. 1971, *41*, 85-90.

Crewe, J. C. The effect of study strategies on the retention of college text material. *Journal of Reading Behavior*, Spring 1969, *1*, 42-45.

Crouse, J. H. Retroactive interference in reading prose materials. *Journal of Educational Psychology,* Feb. 1971, *62,* 39-44.

Curry, R. L., & Roberts, B. H. A comparison of four practice schedules in college developmental reading classes. In Schick and May (1970b, 65-70).

Dare, L. J. A study of remedial programs in public two-year colleges. (Doctoral dissertation, University of North Carolina at Chapel Hill, 1970.) *Dissertation Abstracts International,* Feb. 1971, *31,* 3896A. Ann Arbor: University Microfilms, Order No. 71-3549.

Defrain, D. M. The effects of self-concept and selected personal and educational variables upon attrition in a non-credit college reading improvement program. (Doctoral dissertation, Oklahoma State University, 1970.) *Dissertation Abstracts International,* Apr. 1971, *31,* 5195A. Ann Arbor: University Microfilms, Order No. 71-11, 131.

Dinnan, J. A., Bickley, R. T., & Cowart, H. Syntagmatic oral responses and SAT scores of college freshmen. In Greene, (1971, 17-19).

DiSalvi, R. D. An investigation of the relationship between a study techniques and and reading improvement program and the academic performance of undergraduate evening students. (Doctoral dissertation, University of Pennsylvania, 1970.) *Dissertation Abstracts International,* Nov. 1970, *31,* 2174A. Ann Arbor: University Microfilms, Order No. 70-22, 871.

DiSalvi, R. D. A remedial programs for undergraduate evening students. *Adult Education,* Spring 1971, *21,* 189-195.

Dulin, K. L., & Otto, W. Backgrounds, felt educational needs, and expectations of adults seeking advanced training in reading. In Schick and May (1970a, 116-121).

Dwyer, F. M. Effect of questions on visualized instruction. *Journal of Psychology,* July 1971, *78,* 181-183.

Dwyer, F. M. Questions as advanced organizers in visualized instruction. *Journal of Psychology,* July 1971, *78,* 261-264.

Dwyer, F. M. Student perceptions of the instructional effectiveness of black and white and colored illustrations. *Journal of Experimental Education,* Fall 1971, *40,* 28-34.

Ewing, J. B. The role of the skills instructor in the college reading improvement program. In Schick and May (1970b, 71-75).

Fant, B. J. The "talking typewriter" in a program for the adult disadvantaged. (Doctoral dissertation, University of Florida, 1969.) *Dissertation Abstracts International,* Nov. 1970, *31,* 2000A. Ann Arbor: Universtiy Microfilms, Order No. 70-20, 751.

Farr, R., & Smith, C. B. The effects of test item validity on total test reliability and validity. In Schick and May (1970a, 122-134).

Florak, E. L. Changes in selected variables resulting from group counseling in a developmental reading course. (Doctoral dissertation, University of Pittsburgh, 1970.) *Dissertation Abstracts International,* Apr. 1971, *31,* 5123A. Ann Arbor: University Microfilms, Order No. 71-8773.

Follman, J., Lowe, A. J., & Miller W. Typeface and multiple option format. In Schick and May (1970a, 135-140).

Follman, J., Lowe, A. J., & Miller, W. Graphics variables and reliability and level of essay grages. *American Educational Research Journal,* Mar. 1971, *8,* 365-374.

Follman, J., Miller, W., & Burg, E. Statistical analysis of three critical thinking tests. *Educational Psychological Measurement,* Summer 1971, *31,* 519-520.

Forster, K. I. Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics,* Oct. 1970, *8,* 215-221.

Foss, D. J. Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* Dec. 1970, *9,* 699-706.

Foss, D. J., & Cairns, H. S. Some effects of memory limitation upon sentence comprehension and recall. *Journal of Verbal Learning and Verbal Behavior,* Oct. 1970, *9,* 541-547.

Franklin, T. J. Survey of diagnostic procedures in university and college reading clinics. (Doctoral dissertation, Oklahoma State University, 1969.) *Dissertation Abstracts International,* Feb. 1971, *31,* 3774A. Ann Arbor: University Microfilms, Order No. 70-21, 391.

Frase, L. T., & Silbiger, F. Some adaptive consequences of searching for information in a text. *American Eduactional Research Journal,* Nov. 1970, *7,* 553-560.

Funkhouser, G. R., & Maccoby, N. Communicating special science information to a lay audience. *Journal of Communication*, Mar. 1971, *21*, 58-71.

Furbee, J. W. Writing complexity of college freshmen. (Doctoral dissertation, West Virginia University, 1970.) *Dissertation Abstracts International*, June 1971, *31*, 6275A-6276A. Ann Arbor: University Microfilms, Order No. 71-15, 470.

Gaertner, S. L., & Seidenberg, B. Familiarity and value effects upon stereoscopic word recognition. *Perceptual and Motor Skills*, June 1971, *32*, 703-708.

Garofalo, V. J. A college reading program. *Journal of the Reading Specialist*, Dec. 1970, *10*, 108-111.

Geeslin, R. H. The field test of a material for teaching reading and the general knowledge area of adult basic education: an experimental and descriptive study. (Doctoral dissertation, Florida State University, 1969.) *Dissertation Abstracts International*, Sept. 1970, *31*, 998A. Ann Arbor: University Microfilms, Order No. 70-16, 327.

Geizer, R. Advertising in Ebony: 1960 and 1969. *Journalism Quarterly*, Spring 1970, *48*, 131-134.

Gelso, C. J., Osterhouse, R., & Bodden, J. L. Nonintellective factors in improvement during an educational skills course. *Journal of Counseling Psychology*, Sept. 1971, *18*, 503-504.

Geyer, J. J. A new eye movement measure of reading efficiency. In Schick and May (1970a, 168-171).

Gibson, W. D. Relationship between difficulty levels of assigned English texts and reading ability of community college students. (Doctoral dissertations, University of Southern California, 1971.) *Dissertation Abstracts International*, June 1971, *31*, 6362A-6363A. Ann Arbor: University Microfilms, Order No. 71-16, 411.

Greene, F. P. (Ed.), *Reading: the right to participate*. National Reading Conference 20th Yearbook. Milwaukee: the Conference, 1971.

Grotelueschen, A., & McQuarrie, D. Cross-validation of the quick word test as an estimator of adult mental ability. *Adult Education*, Fall 1970, *21*, 14-19.

Gruner, C. R. Ad hominem satire as a persuader: an experiment. *Journalism Quarterly*, Spring 1971, *48*, 128-131.

Guice, B. M. The use of the cloze procedure for improving reading comprehension of college students. *Journal of Reading Behavior*, Summer 1969, *1*, 81-92.

Guthrie, J. T. Feedback and sentence learning. *Journal of Verbal Behavior*, Feb. 1971, *10*, 23-28.

Hakes, D. T., & Foss, D. J. Decision processes during sentence comprehension: effects of surface structure reconsidered. Perception and Psychophysics, Dec. 1970, *8*, 413-416.

Hámilton, H. W. & Deese, J. Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, Apr. 1971, *10*, 163-170.

Hansom, G. R., &Taylor, R. G. Interaction of ability and personality: another look att he dropout problem in an institute of technology. *Journal of Counseling Psychology*, Nov. 1970, *17*, 540-545.

Hardyck, C. D., & Petrinovich, L. F. Subvocal speech and comprehension level as a function of the difficulty level of reading material. *Journal of Verbal Learning and Verbal Behavior*, Dec. 1970, *9*, 647-652.

Hiett, R., et al. A study of the effectiveness of gun control advertising. *Journalism Quarterly*, Autumn 1969, *46*, 542-544.

Holmes, L. B. A comparative study of the academic performance of two groups of entering college freshmen. (Doctoral dissertation, North Texas State University, 1970.) *Dissertation Abstracts International*, Apr. 1971, *31*, 5232A-5233A. Ann Arbor: University Microfilms, Order No. 71-8677.

Howe, M. J. A. Using students' notes to examine the role of the individual learner in acquiring meaningful subject matter. *Journal of Educational Research*, Oct. 1970, *64*, 61-63.

Hultgren, D. D. The role of the individual learning center in effecting educational change. In Schick and May (1970b, 89-94).

Hultgren, D., & Crewe, J. Time as a factor in the retention of college text material. *Journal of Reading Behavior*, Fall 1969, *1*, 49-52.

Ilika, J. The extent of teacher preparation and the development of word analysis knowledge among teacher education students. *Journal of the Reading Specialist*, Mar. 1971, *10*, 124-127.

Ilika, J. The third annual report of the phonetic skills of teachers. In Schick and May (1970b, 95-105).

Jamieson, B. D. Attitude, plausibility, and the learning and recall of controversial material. *Journal of Psychology*, Nov. 1970, *76*, 169-174.

Johnson, D. C. The effect of reading and punctuation variations upon reading comprehension. In Greene (1971, 304-311).

Johnson, J. C., II, Jacobson, M. D., & Beaber, J. D. An analysis of reading proficiency and its improvement in university-accepted rural secondary students. In Schick and May (1970a, 193-197).

Kanungo, R. N., & Dutta, S. Measure of communicative meaning of words: the semantic concensus index. *Psychological Reports*, Aug. 1971, *219*, 295-306.

Kaplan, I. T., & Cawellas, T. Effect of context on verbal recall. *Journal of Verbal Learning and Verbal Behavior*, Apr. 1971, *10*, 207-212.

Kassarjian, H. H. The Negro and American advertising, 1946-1965. *Journal of Marketing Research*, 1969, *6*, 29-39.

Keenen, C. B. Comparative achievement of two substandard groups of university freshmen. *Journal of Educational Research*, Jan. 1971, *64*, 203-204.

Kerstiens, G. Open enrollment: challenge to the reading/study skills program. In Schick and May (1970b, 106-112).

Kline, F. G. Media time budgeting as a function of demographics and life style. *Journalism Quarterly*, Summer 1971, *48*, 211-221.

Knight, D., & Alcom, J. D. Comparison of the performance of educationally disadvantaged adults and elementary children on selected measures of reading performance. In Schick and May (1970b, 113-117).

Lawrence, D. H. Temporal numerosity estimates for word lists. *Perception and Psychophysics*, Aug. 1971, *10*, 75-78. (a)

Lawrence, D. H. Two studies of visual search for word targets with controlled rates of presentation. *Perception and Psychophysics*, Aug. 1971, *10*, 85-89. (b)

Leslie, R., & Calfee, R. C. Visual search through word lists as a function of grade level, reading ability, and target repetition. *Perception and Psychophysics*, Sept. 1971, *10*, 169-171.

Little, S. W. What is black and white and the tenth largest industry in the country? *Saturday Review*, Feb. 13, 1971, *54*, 54-55.

Locke, J. L. Phonemic processing in silent reading. *Perceptual and Motor Skills*, June 1971, *32*, 905-906.

Loewenthal, K. A study of imperfectly acquired vocabulary. *British Journal of Psychology*, May 1971, *62*, 225-233.

Lowe, A. J., & Stefurak, D. W. The college reading improvement programs of Georgia: 1969-70. In Schick and May (1970b, 118-124).

Lynch, M. D. An index of syntactic dispersion. In Schick and May (1970a, 213-220).

MacGregor, N. The effect of the location of questions in reading material on long-term retention of specific facts by adult subjects. (Doctoral dissertation, Indiana University, 1970.) *Dissertation Abstracts International*, April 1971, *31*, 5110A. Ann Arbor: University Microfilms, Order No. 71-6874.

Mallory, V. C. Effectiveness of practice in extracting and processing information under supervised or non-supervised reading instruction. (Doctoral dissertation, University of Alabama, 1970.) *Dissertation Abstracts International*, Nov. 1970, *31*, 2188A. Ann Arbor: University Microfilms, Order No. 71-17, 942.

Martinez, A. J. Analysis of the present status of the teaching of English as a second language to Puerto Rican adults in New York City. (Doctoral dissertation, New York University, 1970.) *Dissertation Abstracts International*, May 1971, *31*, 5749. Ann Arbor: University Microfilms, Order No. 71-13, 619.

Maxwell, M. J. Effects of practice and learning strategies on speed of scanning for phrases in meaningful material. In Schick and May (1970a, 226-232).

May, R. B., & Tryk, H. E. Word sequence, word frequency, and free recall. *Canadian Journal of Psychology*, Oct. 1970, *24*, 299-304.

McClaskey, H. C. Bibliotherapy with emotionally disturbed patients: an experimental study. (Doctoral dissertation, University of Washington, 1970.) *Dissertation*

*Abstracts International,* Apr. 1971, *31,* 5205A-5206A. Ann Arbor: University Microfilms, Order No. 71-8523.

McClellan, D. A. Reading ability of junior college students and readability of assigned texts. In Greene (1971, 348-354).

McCroskey, J. C., & Wright, D. W. A comparison of the effects of punishment-oriented and reward-oriented messages in persuasive communication. *Journal of Communication,* Mar. 1971, *12,* 82-93.

McDaniel, J. W. Neighborhood youth corps goes to community college. *Junior College Journal,* Apr. 1971, *41,* 14-17.

McEwen, W. J., & Greenberg, B. S. The effects of message intensity on receiver evaluations of source, message, and topic. *Journal of Communication,* Dec. 1970, *20,* 340-350.

McLeod, J., & Anderson, J. An approach to assessment of reading ability through information transmission. *Journal of Reading Behavior,* Spring 1970, *2,* 116-143.

Meehan, Sr. M. T. The effects of instruction based on elements of critical reading upon the questioning patterns of preservice teachers. (Doctoral dissertation, Indiana University, 1970.) *Dissertation Abstracts International,* Jan. 1971,*31,* 3176A. Ann Arbors: University Microfilms, Order No. 70-26, 940.

Meyer, T. P., & Miller, W. C. Emphasis and non-emphasis radio newscast delivery. *Journalism Quarterly,* Spring 1970, *47,* 144-147.

Mickleson, N. I. Meaningful indices for 120 mass and count nouns for children aged 14 years and for university students. *Journal of Educational Research,* Sept. 1971, *65,* 27-30.

Miller, T. K. Artenen, W. W., & Duncan, J. A. High-risk students in junior college: a special program. *Personnel and Guidance Journal,* Mar. 1971, *47,* 537-543.

Mueller, T. H. Immediate audio and visual confirmation: "breakthrough" for the low-aptitude language student. *Audio-Visual Communication Review,* Winter, 1969, *17,* 464-465.

Mujahid, S. al Coverage of Pakistan in three U. S. newsmagazines. *Journalism Quarterly,* Spring 1970, *47,* 126-130.

Mustico, T. W. Some implications from paired associate learning on the development of reading readiness. In Greene (1971, 283-291).

Newman, H. How to prevent an open door from becoming a revolving one: in-service training in reading and study skills for city university faculty. *Journal of the Reading Specialist,* May 1971, *10,* 234-245.

Newsome, R. S., & Gaite, J. H. Prose learning: effects of pretesting and reduction of passage length. *Psychological Reports,* Feb. 1971, *28,* 123-129.

Opison, S. H. Student appraisal of a preregistration college counseling program. *Journal of College Student Personnel,* Jan. 1970, *10,* 47-51.

Otto, W., & Dulin, K. L. How to succeed in reading by really trying. In Schick and May (1970b, 134-141).

Pace, R. A. The role of auditory-visual equivalences and the applicability of automated programmed learning procedure in teaching severely retarded adults to read. Doctoral dissertation, University of Wisconsin) *Dissertation Abstracts International,* June 1971, *31,* 7607B. Ann Arbor: University Microfilms, 1971. Order No. 71-9346.

Paivio, A. Imagery and deep structure in the recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior,* Feb. 1971, *10,* 1-12.

Paivio, A., & O'Neill, B. J. Visual recognition thresholds and dimension of word meaning. *Perception and Psychophysics,* Nov. 1970, *8,* 273-275.

Palmatier, R. A. The last 2 r's: a research view. In Wark (1971, 120-130).

Parent, P. P. Toward the nucleation stage: an investigation of the effectiveness of three methods of teaching reading comprehension in a college course in intermediate French. (Doctoral dissertation, Pennsylvania State University, 1969.) *Dissertation Abstracts International,* Aug. 1970, *31,* 683A. Ann Arbor: University Microfilms, Order No. 70-13, 820.

Payne, D. B. The effect of participation in a college reading improvement program on retention of reading skills and academic achievement. (Doctoral dissertation, Northwestern State University of Louisiana, 1971.) *Dissertation Abstracts International,* June 1971, *31,* 6286A-6287A. Ann Arbor: University Microfilms, Order No. 71-16, 257.

Pearson, D. T. The effects of a combined reading and group counseling program on community college students. (Doctoral dissertation, University of Montana, 1970.) *Dissertation Abstracts International*, Dec. 1970, *31*, 2690A. Ann Arbor: University Microfilms, Order No. 70-24, 131.

Pepper, R. S. The study skills and academic achievement of marginal admission students. In Schick and May (1970a, 248-253).

Pepper, R. S., & Drexler, J. A., Jr. Attitudinal factors among marginal admission students. In Wark (1971, 62-88).

Phi Delta Kappan. Functional illiteracy reported high. *Phi Delta Kappan*, Dec. 1970, *52*, 253.

Phillips, G. O., Sr. The relative effectiveness of three instructional approaches upon the reading, study habits and attitudes, and academic performances of disadvantaged black college freshmen. (Doctoral dissertation, University of Pennsylvania, 1969.) *Dissertation Abstracts International*, Sept. 1970, *31*, 1084A. Ann Arbor: University Microfilms, Order No. 70-15, 460.

Phillips, G. O., Sr. Study habits and attitudes of disadvantaged students in a college reading and study skills program. In Schick and May (1970b, 152-157).

Railsback, L. L. A V.I.T.A.L.I.T.Y. center: supplementary reading materials and services. In Schick and May (1970b, 158-161).

Rankin, E. F., & Hess, A. K. The measurement of internal (intra-article) reading flexibility. In Schick and May (1970a, 254-262).

Rentel, V. M., & Kennedy, J. J. Word length and form class factors in a sample of continuous prose. In Schick and May (1970a, 263-269).

Reynolds, R. J., & Lundrum, W. Step-wise analysis of word association data. In Schick and May (1970a, 270-276).

Reynolds, R. J., Palmatier, R. A., & Ulmer, C. Effect of mode of input on ideational fluency in adult training groups. *Journal of Reading Behavior*, Spring 1969, *1*, 53-64.

Rogers, R. Education and political involvement in USSR elite newspaper reading. *Journalism Quarterly*, Winter 1970, *47*, 735-745.

Ross, L. Forecasting the academic achievement of engineering freshmen. *Journal of Educational Research*, Mar. 1971, *64*, 307-310.

Rossiter, C. M., Jr. Chronological age and listening of adult students. *Adult Education*, Fall 1970, *21*, 40-43.

Rossiter, C. M., Jr. Rate of presentation effects on recall of facts and of ideas and on generation of inferences. *AV Communication Review*, Fall 1971, *19*, 313-324.

Sachsman, D. B. A test of "loading": new measure of bias. *Journalism Quarterly*, Winter 1970, *47*, 759-762.

Sainty, G. E. Predicting drop-outs in adult education courses. *Adult Education*, Summer 1971, *21*, 223-230.

Samuels, S. J., & Chen, C. C. Comparison of word recognition strategies of adults and children. In Greene (1971, 73-77).

Santiago, R. L. A contrastive analysis of some rhetorical aspects in the writing in Spanish and English of Spanish-speaking college students in Puerto Rico. (Doctoral dissertation, Columbia University, 1970.) *Dissertation Abstracts International*, June 1971, *31*, 6368A-6369A. Ann Arbor: University Microfilms, Order No. 71-14, 334.

Schick, G. B., & May, M. M. (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook. (Vol. I) Milwaukee: the Conference, 1970. (a)

Schick, G. B., & May, M. M. (Eds.), *Reading: process and pedagogy*. National Reading Conference 19th Yearbook. (Vol. II) Milwaukee: the Conference, 1970. (b)

Schneider, V. L. A study of the effectiveness of emphasizing the teaching of reading skills to improve composition skills in remedial English classes at Kansas City Community Junior College. (Doctoral dissertation, University of Kansas, 1970.) *Dissertation Abstracts International*, June 1971, *31*, 6369A-6370A. Ann Arbor: University Microfilms, Order No. 71-13, 359.

Singer, C. R. Eye movements, reading comprehension and vocabulary as effected by varying visual and auditory meodalities in college students. (Doctoral dissertation, New Mexico State University, 1970.) *Dissertation Abstracts International*, July 1970, *31*, 164A. Ann Arbor: University Microfilms, Order No. 70-11, 514.

Smith, J. L. Student perceptions of a reading and study skills center. (Doctoral dissertation, Indiana University, 1969.) *Dissertation Abstracts International*, July 1970, *31*, 181A. Ann Arbor: University Microfilms, Order No. 70-12, 953.

Smith, J. M., & McCombs, M. E. The graphics of prose. *Journalism Quarterly*, Spring 1971, *48*, 134-136.

Smith, K. V., Schremser, R. & Putz, V. Binocular coordination in reading. *Journal of Applied Psychology*, June 1971, *55*, 251-258.

Smith, R. J., Hesse, K., & Otto, W. Factors influencing junior high school teacher's attitudes toward teaching reading in the content areas. In Schick and May (1970b, 177-184).

Smith, W. A. A study of the personality assessment system and its utility in the prediction of performance in a reading improvement course. (Doctoral dissertation, University of Missouri—Columbia, 1970.) *Dissertation Abstracts International*, Apr. 1971, *31*, 5216A. Ann Arbor: University Microfilms, Order No. 71-8391.

Sommer, R. Study conditions in student residences. *Journal of College Student Personnel*, July 1969, *10*, 270-274.

Spector, N. J. The impact of the editorial page on a municipal referendum. *Journal Quarterly*, Winter 1970, *47*, 762-766.

Stempel, G. H., III. Visibility of blacks in news and news-picture magazines. *Journalism Quarterly*, Summer 1971, *48*, 337-339.

Stewart, E. W. Reading improvement program for college students. In Schick and May (1970b, 202-207).

Sticht, T. G. Failure to increase learning using the time saved by the time compression of speech. *Journal of Educational Psychology*, Feb. 1971, *62*, 55-59.

Stieglitz, F. B. The effect of sentence length and grammatical structure on repeatability of sentences by native and foreign speakers of English. (Doctoral dissertation, Columbia University, 1970.) *Dissertation Abstracts International*, June 1971, *31*, 6417A. Ann Arbor: University Microfilms, Order No. 71-14,336.

Swift, P. Keeping up . . . with youth. *Parade*, Aug. 1, 1971, *10*.

Taschow, H. G. Reading gains measured by two different methods. *Saskatchewan Journal of Educational Research and Development*, Spring 1971, *1*, 38-44.

Tichenor, P. J., Olien, C. N., Harrison, A., & Donohue, G. Mass communciation systems and communication accuracy in science news reporting. *Journalism Quarterly*, Winter 1970, *47*, 673-683.

Tobias, S., & Abramson, T. Interaction among anxiety, stress, response mode and familiarity of subject matter on achievement from programmed instruction. *Journal of Educational Psychology*, Aug. 1971, *62*, 357-364.

Todd, W. B., & Kessler, C. C., III. Influence of response mode, sex, reading ability, and level of difficulty on four measures of recall of meaningful written material. *Journal of Educational Psychology*, June 1971, *62*, 229-234.

Tuinman, J., & Hafner, L. E. Information value of adjectives in relation to their position in sentences. *Psychological Reports*, June 1971, *28*, 987-990.

Vraa, C. W. Predicting academic achievement of Canadian college freshmen. *Journal of College Student Personnel*, July 1971, *12*, 303-308.

Wali, A. J. T. Adult education in Iraq: a comparative study of the verbal method and audio-visual techniques int he learning process. (Doctoral dissertation, University of California, Los Angeles, 1970.) *Dissertation Abstracts International*, Sept. 1970, *31*, 1001A. Ann Arbor: University Microfilms, Order No. 70-16, 339.

Wark, D. M. (Ed.), *College and adult reading*. North Central Reading Association 6th Yearbook. St. Paul: the Association, 1971.

Wark, D. M. Test panic, daydreaming, and procrastination. In Wark (1971, 131-151).

Warren, R. M., & Obusek, C. J. Speech perception and phonemic restorations. *Perception and Psychophysics*, Mar. 1971, *9*, 359-362.

Warsh, H. E. Behavior modification of adult illiterates and functional illiterates who elarned to read. (Doctoral dissertation, Wayne State University, 1969.) *Dissertation Abstracts International*, Sept. 1970, *31*, 1002A. Ann Arbor: University Microfilms, Order No. 70-16, 527.

Watts, M. W., Jr. Differences in educational level and subject students. *Adult Education*, Winter 1971, *21*, 27-36.

Weaver, W. W., Holmes C. C., & Reynolds, R. J. The effect of reading variation and punctuation condition upon reading comprehension. *Journal of Reading Behavior*, Winter 1970, *2*, 75-84.

Weaver, W. W., Kingston, A. J. Bickley, A. C., & White, W. F. Information flow difficulty in relation to reading comprehension. *Journal of Reading Behavior*, Summer 1969, *1*, 41-49.

Weigel, R. G. Laterality, mirror-writing and mirror-reading. *Perceptual and Motor Skills*, June 1971, *32*, 969-970.

Weigel, R. G., Weigel, V. M., & Herbert, J. A. Non-volunteer subjects: temporal effects. *Psychological Reports*, Feb. 1971, *28*, 191-192.

Whitehill, R. P., & Rubin, S. J. Effectiveness of instrumental and traditional methods of college reading instruction. *Journal of Experimental Education*, Spring 1971, *39*, 85-88.

Williams, R. L. On quantifying and measuring the reading backgrounds of students just entering college. (Doctoral dissertation, University of Pittsburg, 1969.) *Dissertation Abstracts International*, Sept. 1970, *31*, 1156A. Ann Arbor: University Microfilms, Order No. 70-13, 791.

Williams, R. P. Predicting comprehension gains of college freshmen. *Journal of Reading Behavior*, Spring 1970, *2*, 166-172.

Yang, S. A readability formula for Chinese language. (Doctoral dissertation, University of Wisconsin, 1971.) *Dissertation Abstracts International*, June 1971, *31*, 6536A. Ann Arbor: University Microfilms, Order No. 71-12, 723.

Yarington, D. J. A review of the research literature dealing with high school and college students' time spent reading and studying. *Journal of the Reading Specialist*, Dec. 1970, *10*, 96-107.